HOKKAIDO UNIVERSITY

Title	二方向の固定効果を持つパネルデータ標本選別モデルにおける二段階推定量のバイアス修正
Author(s)	高木, 真吾; 高良, 佑樹
Citation	經濟學研究, 67(2), 5-15
Issue Date	2017-12-19
Doc URL	http://hdl.handle.net/2115/68051
Type	bulletin (article)
File Information	3ES_67(2)_005.pdf

Instructions for use

Hokkaido University Collection of Scholarly and Academic Papers : HUSCAP

経済学研究 67-2
北海道大学 2017.12

二方向の固定効果を持つパネルデータ標本選別モデルにおける二段階推定量のバイアス修正

高木真吾* ・ 高良佑樹

1．はじめに

本稿では，パネルデータを用いて，二方向の固定効果をもつ二段階推定法を用いた推定量のバイアス修正，特に標本選別モデルの推定を行う際のバイアス修正について議論する。標本サイズを，横断面方向に N，時系列方向に T という形で表記し，N と T が同程度の速度で増加するタイプの漸近理論を考える。こうした標本に関する前提は，経済学においては，証券市場の分析（Bai［2009］）や国別のマクロデータを用いた分析（Maddala and Wu［1999］）に用いられるとともに，二者間データ（Dyadic Data）の経済学分野での代表例であるネットワークデータや貿易データを用いた実証研究においても用いられている（Cameron and Miller［2014］）。特に貿易の内延（intensive margin）と外延（extensive margin）に焦点を当てた Helpman, Melitz, and Rubinstein［2008］では，貿易関係に入るか否かという選択と貿易量の決定という構造を持つ標本選別モデルのようなモデルを用いて，貿易理論に関する構造分析を行っている。

Helpman, Melitz, and Rubinstein［2008］では比較的多くの国を用いているため，N と T が十分大きいという前提の下で非線形パネルデータモデルの推定を行っていると考えられる。それゆえ付随パラメータ（incidental parameters）の存在に伴い推定量にバイアスが生じる問題（Neyman and Scott［1948］）を回避し，理論的には一致性を持つことを示すことができるものの，その有限標本における推定量の振る舞いについては未解明である。加えて，近年の非線形パネルデータモデルに関する理論的研究が明らかにしたのは，N と T が同程度の速度で増加する漸近理論を考える限り，推定量の漸近分布の平均には無視できないバイアスが生じうることである。

つまり推定量にオーダー N^{-1} とオーダー T^{-1} のバイアスが存在するとき，推定量の漸近分散は $\sqrt{NT}$ で基準化されるが，このとき上記の二つのバイアス項は，それぞれ $\sqrt{NT}/T=\sqrt{N/T}$，$\sqrt{NT}/N=\sqrt{T/N}$ という次数のバイアスとなるため，N と T が同程度で増加するとき，この二つのバイアス項は漸近的にも消滅しない。本稿では，標本選別モデルの二段階推定法に関して，時系列方向および横断面方向に生じるオーダー N^{-1} とオーダー T^{-1} のバイアスを修正する方法を提示し，その有限標本での振る舞いを確認する。

2. モデルと推定量

二段階推定において，推定に用いる目的関数が以下のように与えられるものとする。固定効果パ

* この研究は JSPS 科研費 JP15H03358 の助成を受けたものである。

ラメータをまとめて，$\phi_j = (\boldsymbol{\alpha}'_{j,1}, \boldsymbol{\alpha}'_{j,2}, \ldots, \boldsymbol{\alpha}'_{j,N}, \boldsymbol{\gamma}'_{j,1}, \boldsymbol{\gamma}'_{j,2}, \ldots, \boldsymbol{\gamma}'_{j,T})'$（$j = 1,2$）と表記し，

$$Q_{1,NT}(\boldsymbol{\theta}_1, \phi_1) = \sum_{i=1}^{N}\sum_{t=1}^{T} q_{1,it}(\boldsymbol{\theta}_1, \underbrace{\boldsymbol{\alpha}_{1,i} + \boldsymbol{\gamma}_{1,t}}_{\boldsymbol{\pi}_{1,it}:M_1\times 1}), \qquad \boldsymbol{\eta}_1 = (\underbrace{\boldsymbol{\theta}'_1}_{1\times K_1}, \underbrace{\boldsymbol{\alpha}'_1}_{1\times NM_1}, \underbrace{\boldsymbol{\gamma}'_1}_{1\times TM_1})' \tag{1}$$

$$Q_{2,NT}(\boldsymbol{\theta}_2, \phi_2; \boldsymbol{\lambda}(\boldsymbol{\eta}_1)) = \sum_{i=1}^{N}\sum_{t=1}^{T} q_{2,it}(\boldsymbol{\theta}_2, \underbrace{\boldsymbol{\alpha}_{2,i} + \boldsymbol{\gamma}_{2,t}}_{\boldsymbol{\pi}_{2,it}:M_2\times 1}; \underbrace{\boldsymbol{\lambda}_{it}(\boldsymbol{\eta}_1)}_{L\times 1}), \ \boldsymbol{\eta}_2 = (\underbrace{\boldsymbol{\theta}'_2}_{1\times K_2}, \underbrace{\boldsymbol{\alpha}'_2}_{1\times NM_2}, \underbrace{\boldsymbol{\gamma}'_2}_{1\times TM_2})', \tag{2}$$

ただし，$\boldsymbol{\lambda}_{it}(\boldsymbol{\eta}_1)$ は第一段階の推定結果から得られる結果を用いたコントロール関数であり，通常の標本選別モデル（例えば，Heckman［1976］，Wooldridge［2010］）を考えると，

$$d_{it} = 1\left\{\mathbf{x}'_{1,it}\boldsymbol{\delta} + \alpha_{1,i} + \gamma_{1,t} + u_{1,it} \geq 0\right\}, u_{1,it} \sim N(0,1), \ i = 1,2,\ldots,N, t = 1,2,\ldots,T.$$
$$y^*_{it} = \mathbf{x}'_{2,it}\boldsymbol{\beta} + \alpha_{2,i} + \gamma_{2,t} + u_{2,it}, u_{2,it} \sim N(0,\sigma^2), \mathbb{E}[u_{1,it}u_{2,it}] = \rho\cdot\sigma$$

という二本の方程式からなり，前者は選別方程式（selection equation），後者は結果方程式（outcome equation）と呼ばれる。

また結果方程式における従属変数は $d_{it} = 1$ という条件の下でのみ観測されるという観測ルールも適用される。このとき，推定のための目的関数は，

$$q_{1,it} = d_{it}\cdot\log\Phi(\mathbf{x}'_{1,it}\boldsymbol{\delta} + \alpha_{1,i} + \gamma_{1,t}) + (1 - d_{it})\cdot\log\Phi(-\mathbf{x}'_{1,it}\boldsymbol{\delta} - \alpha_{1,i} - \gamma_{1,t})$$
$$q_{2,it} = -\frac{1}{2}(y_{it} - \mathbf{x}'_{2,it}\boldsymbol{\beta} - \zeta\cdot\lambda(\mathbf{x}'_{1,it}\boldsymbol{\delta} + \alpha_{1,i} + \gamma_{1,t}) - \alpha_{2,i} - \gamma_{2,t})^2.$$

ただし，$\zeta = \rho\sigma, \lambda(z) = \phi(z)/\Phi(z)$ であり，ϕ および Φ はそれぞれ標準正規分布の密度関数と分布関数である。

ここで $\boldsymbol{\theta}_1 = \boldsymbol{\delta}, \boldsymbol{\theta}_2 = (\boldsymbol{\beta}, \zeta)$ とし，L=M_1=M_2=1 とすれば，上記のモデル（1），（2）に適合している。その他に上記の枠組みに適合するモデルの例としては Fernández-Val and Vella［2011］や Helpman, Melitz, and Rubinstein［2008］などがあり[1)]，Takara［2017］では Helpman らの方法を文化財貿易の文脈で，L=7，M_1=M_2=1 として応用した。

一般的に第一段階，第二段階の推定量を定義すると，第一段階では

$$\hat{\phi}_1(\boldsymbol{\theta}_1) = \arg\max_{\phi_1} Q_{1,NT}(\boldsymbol{\theta}_1, \phi_1)$$
$$\hat{\boldsymbol{\theta}}_1 = \arg\max_{\theta_1} Q_{1,NT}(\boldsymbol{\theta}_1, \hat{\phi}_1(\boldsymbol{\theta}_1))$$

1）このモデルの拡張例として，交差固定効果，あるいは加法的固定効果と交差固定効果の双方を含むモデルを考えることも可能である。交差固定効果は因子分析との関連が深く，推定結果の解釈に困難を伴うこともあるが，パネル分析の文脈では空間的相関を捉える要因としても用いられることがある（例えば Chudik and Pesaran[2015]）。Chen，Fernández-Val，and Weidner [2014] では交差固定効果を含むモデルでのバイアス修正を提案しているが，そこでのバイアス修正法は二段階推定の文脈にも容易に拡張することが可能であり，加法的固定効果と交差固定効果の双方を含むモデルでの固定効果パラメータの識別については Bai [2009] の方法を利用することができる。

また$\boldsymbol{\theta}_1$の推定量を$\tilde{\boldsymbol{\theta}}_1$として，$\tilde{\phi}_1 = \hat{\phi}(\tilde{\boldsymbol{\theta}}_1)$および$\tilde{\boldsymbol{\eta}}_1 = (\tilde{\boldsymbol{\theta}}_1', \tilde{\phi}_1')'$とおくと，第二段階の推定量は

$$\hat{\phi}(\boldsymbol{\theta}_2) = \arg\max_{\phi_2} Q_{2,NT}(\boldsymbol{\theta}_2, \phi_2; \boldsymbol{\lambda}(\tilde{\boldsymbol{\eta}}_1))$$
$$\hat{\boldsymbol{\theta}}_2 = \arg\max_{\theta_2} Q_{2,NT}(\boldsymbol{\theta}_2, \hat{\phi}_2(\boldsymbol{\theta}_2); \boldsymbol{\lambda}(\tilde{\boldsymbol{\eta}}_1))$$

と定義することができる。

推定に際して，固定効果をすべて同時に推定することは厳密な多重共線性のため不可能である。そこでそれぞれの方向の固定効果が合計としては等しくなる（$\Sigma_{i=1}^{N}\boldsymbol{\alpha}_{j,it} = \Sigma_{t=1}^{T}\boldsymbol{\gamma}_{j,it}$）という条件を課すことで$N+T-1$個のパラメータを独立に推定することができる。つまり上記の線形制約を制約条件とみなした制約条件付き最大化によって推定値を求める方法が Fernández-Val and Weidner［2016］によって提案されている。同値な方法として，αとγのそれぞれ一個ずつを削除し，定数項を加えた形で合計（$N-1$）+（$T-1$）+ 1 = $N+T-1$個の固定効果パラメータを推定することもできる（定数項と制約条件，および（$N-1$）+（$T-1$）個の固定効果推定値ですべての固定効果を復元できる）。こちらは推定手順において，制約条件のない最大化問題を解くことで十分なため，比較的容易に推定値を得ることができる。本稿では後者の方針で推定を行う。

3. 推定量のバイアス修正

Hahn and Newey［2004］では一方向の固定効果を持つ非線形パネルデータモデルの最尤推定量におけるオーダーT^{-1}のバイアスを修正する方法について，明示的にバイアス修正項を考慮する方法とリサンプリング法による方法が議論された。この分野では，複数のバイアス修正法の理論的な比較を行う Hahn and Kuersteiner［2011］，プロビットモデルについて詳細に分析を行った Fernández-Val［2009］，スイッチング回帰モデルの最尤推定のように多変量の方程式体系にも適用可能とすべく，同一方向ではあるが複数の固定効果を含む形へ拡張した Arellano and Hahn［2006］など多くの研究がおこなわれている。

上記の標本選別モデルは，最尤推定を行う限りは Arellano and Hahn［2006］の枠組みで分析可能であるが，推定値，特に誤差項間の相関係数を求める上で困難を伴うため，Heckman［1976］による二段階推定法をパネルデータの場合にも援用することが多い（例えば Helpman, Melitz, and Rubinstein［2008］）。Fernández-Val and Vella［2011］は，上記文献のバイアス修正法を二段階推定法へ適用可能な形に拡張し，一方向の固定効果を持つ非線形パネルデータモデルについて，オーダーT^{-1}のバイアス修正法を提案した。

二方向の固定効果を持つモデルについては，Fernández-Val and Weidner［2016］においてオーダーT^{-1}およびN^{-1}のバイアスを修正する方法が議論されているが，標本選別モデルの推定に適用可能な二段階推定法を含むようなケースについては既存文献においては知られていない。以下では，第二段階の推定量におけるオーダーT^{-1}およびN^{-1}のバイアス項を明示的に示し，その修正推定量の性質を見る。なお，表記ルールに関しては基本的に Fernández-Val and Weidner［2016］と同様の記号を用いており，データ発生過程に関する仮定も Fernández-Val and Weidner［2016］と同様であることを想定している。一致性の証明，バイアス項の厳密な導出過程および必要な仮定などについては稿を改めて提示する。

第一段階のパラメータ$\boldsymbol{\theta}_1$，及び第二段階のパラメータ$\boldsymbol{\theta}_2$のバイアス修正済み推定量に関する一

般公式は，形式的に以下のように与えられる。以下で用いている期待値演算$\mathbb{E}$は，固定効果をあたかも定数のように見なした条件付きの期待値であることに注意する。

$$\tilde{\boldsymbol{\theta}}_1 \equiv \hat{\boldsymbol{\theta}}_1 - \mathcal{W}_1^{-1}\mathbf{b}_1, \tilde{\boldsymbol{\theta}}_2 \equiv \hat{\boldsymbol{\theta}}_2 - \mathcal{W}_2^{-1}\mathbf{b}_2$$

ただし，$\hat{\boldsymbol{\theta}}_1, \hat{\boldsymbol{\theta}}_2$は前節に定義した通りで，$\mathcal{W}_1$と$\mathcal{W}_2$については，

$$\mathcal{W}_1 = \mathbb{E}[Q_{1,\theta\theta} + \Xi_1' Q_{1,\phi\phi}\Xi_1] = \sum_{i=1}^{N}\sum_{t=1}^{T}\mathbb{E}\left[\frac{\partial^2 q_{1,it}}{\partial\boldsymbol{\theta}_1\partial\boldsymbol{\theta}_1'} + (\Xi_{1,i\bullet} + \Xi_{1,\bullet t})'\frac{\partial^2 q_{1,it}}{\partial\boldsymbol{\pi}_{1,it}\partial\boldsymbol{\pi}_{1,it}'}(\Xi_{1,i\bullet} + \Xi_{1,\bullet t})\right]$$

$$\mathcal{W}_2 = \mathbb{E}[Q_{2,\theta\theta} + \Xi_2' Q_{2,\phi\phi}\Xi_2] = \sum_{i=1}^{N}\sum_{t=1}^{T}\mathbb{E}\left[\frac{\partial^2 q_{2,it}}{\partial\boldsymbol{\theta}_2\partial\boldsymbol{\theta}_2'} + (\Xi_{2,i\bullet} + \Xi_{2,\bullet t})'\frac{\partial^2 q_{2,it}}{\partial\boldsymbol{\pi}_{2,it}\partial\boldsymbol{\pi}_{2,it}'}(\Xi_{2,i\bullet} + \Xi_{2,\bullet t})\right]$$

と書くことができる。行列の添え字のパラメータは，そのパラメータに関する微係数を表すものとし，Ξ_jについては，$\Xi_j = \overline{\mathcal{H}}_j^{-1}\mathbb{E}[Q_{j,\phi\theta}]$，$(M_j \cdot N + M_j \cdot T)\times K_j$であり，最初の$N$個のブロックを$\Xi_{j,i\bullet}$（$j=1,2$，$i=1,2,\ldots,N$），残りの$T$個のブロックを$\Xi_{j,\bullet t}$（$j=1,2$，$t=1,2,\ldots,T$）と分割する。$\overline{\mathcal{H}}_1$，$\overline{\mathcal{H}}_2$については，固定効果パラメータに関するヘッセ行列である。

第一段階の推定量におけるオーダーT^{-1}およびN^{-1}のバイアス項は$\mathcal{W}_1^{-1}\mathbf{b}_1$であり，$\mathbf{b}_1$については，$\Psi_1 \equiv Q_{1,\theta} + \Xi_1' Q_{1,\phi}$として

$$\mathbf{b}_1 = \mathbb{E}\left[\tilde{\Psi}_{1,\phi}\cdot\Delta_{1,NT}\right] + \mathbb{E}\left[\Psi_{1,\phi\phi}\right]\cdot\frac{\boldsymbol{\sigma}_{11}}{2} = \sum_{i=1}^{N}\mathbf{b}_{1,i\bullet} + \sum_{t=1}^{T}\mathbf{b}_{1,\bullet t} + o_p\left((NT)^{-1/2}\right) \tag{3}$$

ここで，記号上のティルダは平均からの乖離を表し，$\Delta_{1,NT}$は$\Delta_{1,NT} = \overline{\mathcal{H}}_1^{-1}Q_{1,\phi}$，であり，$\boldsymbol{\sigma}_{11}$については，$\boldsymbol{\sigma}_{11} \equiv \mathbb{E}\left[\Delta_{1,NT}\otimes\Delta_{1,NT}\right]$とする。$\mathbf{b}_{1,i\bullet}$および$\mathbf{b}_{1,\bullet t}$の具体的な表現は補論に明記されている。

次に第二段階の推定量におけるオーダーT^{-1}およびN^{-1}のバイアス項は$\mathcal{W}_2^{-1}\mathbf{b}_2$であるが，$\mathbf{b}_2$については以下のように与えられる。$\Psi_2 \equiv Q_{2,\theta} + \Xi_{2,NT}' \cdot Q_{2,\phi}$として，

$$\begin{aligned}\mathbf{b}_2 &= \mathbb{E}\left[\Psi_{2,\lambda}\Lambda_\phi\right]\cdot\boldsymbol{\beta}_1 + \mathbb{E}\left[\tilde{\Psi}_{2,\phi}\cdot\Delta_{2,NT}\right] + \mathbb{E}\left[\left(\tilde{\Psi}_{2,\lambda} + \tilde{\Psi}_{2,\phi}\overline{\mathcal{H}}_2^{-1}\mathbb{E}[Q_{2,\phi\lambda}]\right)\cdot\Lambda_\phi\cdot\Delta_{1,NT}\right] \\ &+ \mathbb{E}\left[\Psi_{2,\phi\phi}\right]\cdot\frac{\boldsymbol{\sigma}_{22}}{2} + \mathbb{E}\left[\Psi_{2,\phi\lambda}(\mathbf{I}_{N+T}\otimes\Lambda_\phi)\right]\cdot\boldsymbol{\sigma}_{21} + \mathbb{E}\left[\Psi_{2,\lambda}\Lambda_{\phi\phi} + \Psi_{2,\lambda\lambda}(\Lambda_\phi\otimes\Lambda_\phi)\right]\cdot\frac{\boldsymbol{\sigma}_{11}}{2} \\ &= \sum_{i=1}^{N}\mathbf{b}_{2,i\bullet} + \sum_{t=1}^{T}\mathbf{b}_{2,\bullet t} + o_p\left((NT)^{-1/2}\right)\end{aligned} \tag{4}$$

ただし，添え字のパラメータは，そのパラメータによる微係数を表しているので，例えば，$Q_{2,\phi\lambda} = \partial^2 Q_{2,NT}/\partial\phi_2\partial\lambda'$は$M_2$（$N+T$）×$LNT$の行列となり，三次の微係数については，$Q_{2,\theta\phi\lambda} = \partial^2 Q_{2,NT}/\partial\theta_2(\partial\phi_2\otimes\partial\lambda)'$と表記することで$K_2$×｛$M_2$（$N$+$T$）・$LNT$｝の行列となる。また$\Lambda_\phi = \partial\boldsymbol{\lambda}/\partial\phi_1'$（$L\times M_1(N+T)$）であり，$\Lambda_{\phi\phi} = \partial^2\boldsymbol{\lambda}/(\partial\phi_1\otimes\partial\phi_1)'$（$L$×$M_1^2$（$N$+$T$）^{2}）である。

$\Delta_{2,NT}$ は $\Delta_{2,NT}=\overline{\mathcal{H}}_2^{-1}\mathcal{Q}_{2,\phi}$ であり，$\boldsymbol{\sigma}_{\ell m}$ については，$\boldsymbol{\sigma}_{\ell m}\equiv\mathbb{E}\left[\Delta_{\ell,NT}\otimes\Delta_{m,NT}\right]$ とする。

最後に，$\mathbf{b}_{2,i\bullet}$，$\mathbf{b}_{2,\bullet t}$ および $\boldsymbol{\beta}_1$ については具体的な表現を補論に示している。

上記の一般的な結果を標本選別モデルの二段階推定に適用する。第一段階のプロビットモデルの推定は Fernández-Val and Weidner［2016］の Example 1 として，バイアス修正項に関する結果が示されている。以下では第二段階のバイアス修正項についてのみ記載する。

補論の一般公式を $L=M_1=M_2=1$ で誤差項が二変量正規分布[2] の場合に適用すると以下のような結果を得る。$\tilde{\mathbf{x}}_{it}$ は，$\mathbf{x}_{it}$ を定数項と（$N-1$）+（$T-1$）個の固定効果部分に回帰した残差であり[3]，

$\tilde{\boldsymbol{\lambda}}_{it,T}=\boldsymbol{\lambda}_{it}-\dfrac{\mathbb{E}_T[d_{is}\lambda_{is}]}{\mathbb{E}_T[d_{is}]}$ ，$\tilde{\boldsymbol{\lambda}}_{it,N}=\boldsymbol{\lambda}_{it}-\dfrac{\mathbb{E}_N[d_{it}\lambda_{it}]}{\mathbb{E}_N[d_{it}]}$ などとして，

$$
\begin{aligned}
\mathbf{b}_{2,i\bullet}=&-\overline{\mathbb{E}}_T\left[\left(d_{it}\cdot\begin{pmatrix}\tilde{\mathbf{x}}_{it}\\ \tilde{\boldsymbol{\lambda}}_{it}\end{pmatrix}\right)\cdot\frac{v_{2,is}}{\mathbb{E}_T[d_{it}]}\right]-\overline{\mathbb{E}}_T\left[\left(d_{it}\cdot\begin{pmatrix}\tilde{\mathbf{x}}_{it}\\ \tilde{\boldsymbol{\lambda}}_{it}\end{pmatrix}\cdot\zeta\cdot\tilde{\boldsymbol{\lambda}}_{it,\pi}\right)\cdot\psi_{1,is}^{(T)}\right]\\
&-\mathbb{E}_T\left[d_{it}\cdot\begin{pmatrix}\tilde{\mathbf{x}}_{it}\\ \tilde{\boldsymbol{\lambda}}_{it}\end{pmatrix}\cdot\left(\tilde{\boldsymbol{\lambda}}_{it,\pi}\boldsymbol{\zeta}\cdot\beta_{1,i\bullet}+\tilde{\boldsymbol{\lambda}}_{it,\pi\pi}\boldsymbol{\zeta}\cdot\frac{\sigma_{11,i\bullet}}{2}\right)\right]-\begin{pmatrix}\mathbf{0}\\ \mathbb{E}_T\left[d_{it}\cdot\tilde{\boldsymbol{\lambda}}_{it,\pi}\tilde{\boldsymbol{\lambda}}_{it,\pi}\cdot\boldsymbol{\zeta}\right]\end{pmatrix}\cdot\sigma_{11,i\bullet}\\
&+\begin{pmatrix}\mathbf{0}\\ \overline{\mathbb{E}}_T\left[\left(d_{it}\cdot\tilde{\boldsymbol{\lambda}}_{it,\pi}v_{2,it}\right)\cdot\psi_{1,is}^{(T)}\right]\end{pmatrix}+\begin{pmatrix}\mathbf{0}\\ \mathbb{E}_T\left[d_{it}\cdot v_{2,it}\cdot\left(\tilde{\boldsymbol{\lambda}}_{it,\pi}\cdot\beta_{1,i\bullet}+\tilde{\boldsymbol{\lambda}}_{it,\pi\pi}\cdot\frac{\sigma_{11,i\bullet}}{2}\right)\right]\end{pmatrix}
\end{aligned}
$$

$$
\begin{aligned}
\mathbf{b}_{2,\bullet t}=&-\mathbb{E}_N\left[\left(d_{it}\cdot\begin{pmatrix}\tilde{\mathbf{x}}_{it}\\ \tilde{\boldsymbol{\lambda}}_{it}\end{pmatrix}\right)\cdot\frac{v_{2,it}}{\mathbb{E}_T[d_{it}]}\right]-\mathbb{E}_N\left[\left(d_{it}\cdot\begin{pmatrix}\tilde{\mathbf{x}}_{it}\\ \tilde{\boldsymbol{\lambda}}_{it}\end{pmatrix}\cdot\boldsymbol{\zeta}\cdot\tilde{\boldsymbol{\lambda}}_{it,\pi}\right)\cdot\psi_{1,it}^{(N)}\right]\\
&-\mathbb{E}_N\left[\left(d_{it}\cdot\begin{pmatrix}\tilde{\mathbf{x}}_{it}\\ \tilde{\boldsymbol{\lambda}}_{it}\end{pmatrix}\cdot\left(\tilde{\boldsymbol{\lambda}}_{it,\pi}\boldsymbol{\zeta}\cdot\beta_{1,\bullet t}+\tilde{\boldsymbol{\lambda}}_{it,\pi\pi}\boldsymbol{\zeta}\cdot\frac{\sigma_{11,\bullet t}}{2}\right)\right)\right]-\begin{pmatrix}\mathbf{0}\\ \mathbb{E}_N\left[d_{it}\cdot\tilde{\boldsymbol{\lambda}}_{it,\pi}\tilde{\boldsymbol{\lambda}}_{it,\pi}\cdot\boldsymbol{\zeta}\right]\end{pmatrix}\cdot\sigma_{11,\bullet t}\\
&+\begin{pmatrix}\mathbf{0}\\ \mathbb{E}_N\left[\left(d_{it}\cdot\tilde{\boldsymbol{\lambda}}_{it,\pi}v_{2,it}\right)\cdot\psi_{1,it}^{(N)}\right]\end{pmatrix}+\begin{pmatrix}\mathbf{0}\\ \mathbb{E}_N\left[d_{it}\cdot v_{2,it}\cdot\left(\tilde{\boldsymbol{\lambda}}_{it,\pi}\cdot\beta_{1,\bullet t}+\tilde{\boldsymbol{\lambda}}_{it,\pi\pi}\cdot\frac{\sigma_{11,\bullet t}}{2}\right)\right]\end{pmatrix}
\end{aligned}
$$

上記において，$\mathbb{E}_T$ は時間軸方向への確率極限を考えており，$\mathbb{E}_N$ は横断面方向への確率極限を考えている。$\overline{\mathbb{E}}_T[f_t g_s]$ は，いわゆる片側長期分散（one-sided long-run variance），$\Sigma_{s=0}^{\infty}\mathrm{cov}(f_t,g_{t-s})$，を意味している。$v_{2,it}$ や $\psi_{1,it}$ は目的関数の一次条件が主たる要素であり，ラグ付き従属変数を含む程度の動学性を持つなら，過去の履歴を所与とした条件付期待値が 0 となり，時間軸方向ではマーチンゲール差分過程となり，過去の値の関数とは無相関になる。そのため $\mathbb{E}[(\Sigma_{t=1}^{T}f_t)\cdot(\Sigma_{t=1}^{T}g_t)]$ において，

2）二段階推定量が一致性を持つ上で，選別方程式・結果方程式に含まれる誤差項の同時分布が正規分布である必要はない（例えば Wooldridge [2010]）。

3）本質的には Ξ_2 が説明変数ベクトルの定数項と（$N-1$）+（$T-1$）個の固定効果部分への直交射影となっていることから従う結果である。

g_tがマーチンゲール差分過程であれば，その確率極限は，両側長期分散ではなく，片側長期分散を考慮するだけで充分である。また，それぞれの方向に含まれる最後の二項は説明変数が外生的であるとき消えることもわかる。$v_{2,it}$と$\psi_{1,it}$については

$$v_{2,it} = y_{it} - \mathbf{x}'_{2,it}\boldsymbol{\beta} - \zeta \cdot \lambda(\mathbf{x}'_{1,it}\boldsymbol{\delta} + \alpha_{1,i} + \gamma_{1,t}) - \alpha_{2,i} - \gamma_{2,t}, \quad \psi^{(T)}_{1,it} = \varphi_{it} \cdot (d_{it} - \Phi_{it}) / \mathbb{E}_T[\varphi_{it}\phi_{it}],$$

$$\psi^{(N)}_{1,it} = \varphi_{it} \cdot (d_{it} - \Phi_{it}) / \mathbb{E}_N[\varphi_{it}\phi_{it}] \text{ であり} (\varphi_{it} \equiv \phi_{it} / \{\Phi_{it}(1-\Phi_{it})\}), \quad \sigma_{11,i\bullet} = \mathbb{E}_T[(\psi^{(T)}_{1,it})^2],$$

$\sigma_{11,\bullet t} = \mathbb{E}_N[(\psi^{(N)}_{1,it})^2]$ とする。以上の結果を用いて第二段階のパラメータθ_2に関するバイアス修正項は

$$\mathcal{B}_2 = \left\{ \sum_{i=1}^{N} \sum_{t=1}^{T} \mathcal{W}_{2,it} \right\}^{-1} \left(\sum_{i=1}^{N} \mathbf{b}_{2,i\bullet} + \sum_{t=1}^{T} \mathbf{b}_{2,\bullet t} \right), \quad \mathcal{W}_{2,it} = -\mathbb{E}\left[d_{it} \cdot \begin{pmatrix} \tilde{\mathbf{x}}_{it} \\ \tilde{\boldsymbol{\lambda}}_{it} \end{pmatrix} \begin{pmatrix} \tilde{\mathbf{x}}_{it} \\ \tilde{\boldsymbol{\lambda}}_{it} \end{pmatrix}' \right] \tag{5}$$

となる。

4. モンテカルロ実験

本節の実験では主に（5）式によるバイアス修正の有限標本特性を確認する。焦点を当てるのは第二段階の推定量であるが，第一段階の推定量の性質が影響を与えることもあり，第一段階の選別方程式（selection equation）に状態依存的要素を含むか否かによって二種類のデータ発生過程を考える。

最初のデータ発生過程は，静学的な選別方程式と結果方程式（outcome equation）を持つモデルである。時点１以降において以下の関係が成立する。

$$\begin{aligned} d_{it} &= 1\{x_{1,it}\delta_1 + x_{2,it}\delta_2 + \alpha_{1,i} + \gamma_{1,t} + u_{1,it} \geq 0\} \\ y^*_{it} &= x_{1,it}\beta_1 + \alpha_{2,i} + \gamma_{2,t} + u_{2,it} \end{aligned} \tag{6}$$

結果方程式の従属変数についての観測ルールは，

$$y_{it} = \begin{cases} y^*_{it} & \text{if } d_{it} = 1 \\ unobserved & \text{if } d_{it} = 0 \end{cases}$$

なので，$d_{it} = 1$ という部分標本に対して以下の回帰モデルを考える。

$$y_{it} = x_{1,it}\beta_1 + \zeta \cdot \lambda(x_{1,it}\delta_1 + x_{2,it}\delta_2 + \alpha_{1,i} + \gamma_{1,t}) + \alpha_{2,i} + \gamma_{2,t} + v_{2,it} \tag{7}$$

固定効果については，互いに独立に$\alpha_{j,i} \sim N(0,1/16)$，$\gamma_{j,t} \sim N(0,1/16)$ とし，説明変数については，

$$x_{j,it} = \frac{1}{2}x_{j,i,t-1} + \alpha_{1,i} + \gamma_{1,t} + v_{j,it}, \; v_{j,it} \sim N(0,3/4), \; j = 1,2, i = 1,2,\ldots,N, t = 1,2,\ldots,T$$

表 1 静学的な選別方程式に基づく標本選別モデル

		選別方程式（6）		結果方程式（7）		選別方程式（6）		結果方程式（7）	
パラメータ		δ_1	δ_2	β	ζ	δ_1	δ_2	β	ζ
真の値		1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0
標本サイズ		（N,T）＝（56,14）				（N,T）＝（100,20）			
修正後	平均	1.008	1.006	0.988	0.939	1.004	1.005	0.989	0.927
（修正前）	（平均）	(1.158)	(1.159)	(0.949)	(0.857)	(1.111)	(1.113)	(0.970)	(0.901)
	標準偏差	0.107	0.105	0.113	0.236	0.065	0.064	0.058	0.131
	（標準偏差）	(0.130)	(0.128)	(0.104)	(0.203)	(0.075)	(0.074)	(0.056)	(0.121)
	RMSE	0.107	0.105	0.114	0.243	0.065	0.064	0.059	0.150
	（RMSE）	(0.205)	(0.204)	(0.116)	(0.249)	(0.134)	(0.136)	(0.063)	(0.157)
d=1 の割合		0.504				0.578			
標本サイズ		（N,T）＝（56,28）				（N,T）＝（100,50）			
修正後	平均	1.005	1.001	0.998	0.987	1.003	1.003	0.997	0.986
（修正前）	（平均）	(1.090)	(1.087)	(0.977)	(0.935)	(1.050)	(1.050)	(0.986)	(0.954)
	標準偏差	0.069	0.065	0.071	0.141	0.039	0.038	0.037	0.076
	（標準偏差）	(0.077)	(0.073)	(0.069)	(0.132)	(0.041)	(0.041)	(0.037)	(0.073)
	RMSE	0.069	0.065	0.071	0.142	0.039	0.038	0.038	0.077
	（RMSE）	(0.119)	(0.114)	(0.073)	(0.147)	(0.065)	(0.064)	(0.040)	(0.087)
d=1 の割合		0.493				0.514			
標本サイズ		（N,T）＝（56,56）				（N,T）＝（100,100）			
修正後	平均	1.000	1.001	1.000	1.004	1.000	1.002	1.000	0.999
（修正前）	（平均）	(1.054)	(1.054)	(0.986)	(0.954)	(1.033)	(1.035)	(0.991)	(0.967)
	標準偏差	0.049	0.049	0.052	0.103	0.026	0.027	0.024	0.054
	（標準偏差）	(0.052)	(0.052)	(0.051)	(0.097)	(0.027)	(0.028)	(0.024)	(0.052)
	RMSE	0.049	0.049	0.052	0.103	0.026	0.027	0.024	0.054
	（RMSE）	(0.075)	(0.075)	(0.053)	(0.108)	(0.043)	(0.045)	(0.025)	(0.061)
d=1 の割合		0.449				0.535			

とする。推定に際しては，固定効果に関する多重共線性を回避するため，$\alpha_{j,1}$ と $\gamma_{j,1}$ を除外し，定数項を加えて，プロビットモデルの推定と逆ミルズ比を加えた線形回帰モデルの推定を行った。標本サイズは，$N = T$ =56 あるいは $N = T$ =100 を上限として，時系列方向へ T =28,14 あるいは T = 50,20 と減少させた場合の影響も見た。実験の反復回数は何れの場合も 1000 回としている。片側長期分散の推定には Newey and West［1987］の方法を用いて非負になることを保証している（モンテカルロ実験の結果表中では，ラグの選択は 5 としたが，T =100 なら 5 より大きくした方がバイアスは小さくなり，T が小さいときはラグ次数も短くした方がバイアスは小さくなる。ただしその程度については小さな影響しかなかった）。

表 1 に静学的な選別方程式に基づく標本選別モデルの実験結果が示されている。修正後の行の結果が（5）によって修正したバイアス修正後の推定値に関する結果で，修正前となっている行は固定効果とともに推定を行っただけのモデルである。理論的には N と T が大きいとき修正前推定量も一致推定量となるが，この実験での標本サイズでは，$N = T$ =56 の場合では，（6）の係数 δ_1，δ_2 に約 5％程度, $N = T$ =100 の場合でも約 3％程度のバイアスが残っている。二段階推定を行った結果方程式については，外生変数の係数 β_1 については 5％以下，コントロール変数の係数 ζ については約 4.6％（$N = T$ =56）から約 3.3％（$N = T$ =100）というバイアスが残っている。標本サイズを時間軸方向へ小さくするとバイアスの程度も大きくなる。

バイアス修正を行った後の結果に注目すると，（7）式の ζ 以外では，著しく時間軸方向と横断面方向の標本サイズ差が大きい場合ですらバイアスの程度は 1％前後であり，修正前推定量よりも小さいバイアスを持つ傾向がある。ζ についても時間軸－横断面の標本サイズ差が大きい場合でも 6.1％から 7.3％の水準のバイアスを持つ程度であり，標本サイズの増加とともにバイアスは解消

表 2　状態依存的な選別方程式に基づく標本選別モデル

		選別方程式（8）		結果方程式（9）		選別方程式（8）		結果方程式（9）	
パラメータ		δ_1	δ_2	β	ζ	δ_1	δ_2	β	ζ
真の値		1.0	0.5	1.0	1.0	1.0	0.5	1.0	1.0
標本サイズ		（N,T）＝（56,14）				（N,T）＝（100,20）			
修正後	平均	1.029	0.423	0.953	0.887	1.020	0.439	0.976	0.915
（修正前）	（平均）	(1.177)	(0.300)	(0.870)	(0.685)	(1.125)	(0.341)	(0.934)	(0.815)
	標準偏差	0.105	0.128	0.180	0.415	0.064	0.081	0.076	0.186
	（標準偏差）	(0.123)	(0.138)	(0.148)	(0.329)	(0.072)	(0.085)	(0.070)	(0.168)
	RMSE	0.109	0.149	0.186	0.430	0.067	0.101	0.080	0.205
	（RMSE）	(0.215)	(0.243)	(0.197)	(0.455)	(0.144)	(0.180)	(0.097)	(0.250)
d=1 の割合		0.581				0.634			
標本サイズ		（N,T）＝（56,28）				（N,T）＝（100,50）			
修正後	平均	1.010	0.471	0.989	0.977	1.006	0.485	0.994	0.981
（修正前）	（平均）	(1.093)	(0.407)	(0.933)	(0.843)	(1.053)	(0.447)	(0.970)	(0.919)
	標準偏差	0.067	0.084	0.125	0.269	0.033	0.049	0.055	0.121
	（標準偏差）	(0.073)	(0.087)	(0.111)	(0.232)	(0.035)	(0.050)	(0.052)	(0.113)
	RMSE	0.067	0.089	0.125	0.270	0.034	0.052	0.055	0.123
	（RMSE）	(0.118)	(0.128)	(0.130)	(0.280)	(0.063)	(0.073)	(0.061)	(0.139)
d=1 の割合		0.562				0.583			
標本サイズ		（N,T）＝（56,56）				（N,T）＝（100,100）			
修正後	平均	1.003	0.486	0.997	1.001	1.002	0.493	0.997	0.997
（修正前）	（平均）	(1.055)	(0.458)	(0.960)	(0.909)	(1.032)	(0.475)	(0.979)	(0.945)
	標準偏差	0.046	0.059	0.087	0.175	0.025	0.035	0.037	0.082
	（標準偏差）	(0.049)	(0.062)	(0.081)	(0.157)	(0.026)	(0.036)	(0.035)	(0.078)
	RMSE	0.046	0.061	0.087	0.175	0.025	0.036	0.037	0.082
	（RMSE）	(0.073)	(0.074)	(0.090)	(0.182)	(0.041)	(0.044)	(0.041)	(0.095)
d=1 の割合		0.535				0.600			

している。係数 ζ はそもそも説明変数とコントロール変数 λ の多重共線性の影響を受けやすく，それゆえ推定が困難であることが知られているが，今回の実験では多重共線性が起きにくいよう制御されているため，良好な結果が得られたと考えられる。

次に，静学的な選別方程式を，状態依存性（state-dependence）を伴う形へ変更したデータ発生過程について考える。

$$\begin{aligned} d_{it} &= 1\{x_{1,it}\delta_1 + d_{i,t-1}\delta_2 + \alpha_{1,i} + \gamma_{1,t} + u_{1,it} \geq 0\} \\ y_{it}^* &= x_{1,it}\beta_1 + \alpha_{2,i} + \gamma_{2,t} + u_{2,it} \end{aligned} \tag{8}$$

結果方程式については，上記と同じ観測ルールを適用すると，

$$y_{it} = x_{1,it}\beta_1 + \zeta \cdot \lambda(x_{1,it}\delta_1 + d_{i,t-1}\delta_2 + \alpha_{1,i} + \gamma_{1,t}) + \alpha_{2,i} + \gamma_{2,t} + v_{2,it} \tag{9}$$

となる。説明変数 x_1 については前出のものと同じで，時点 0 における選別結果は

$$d_{i0} = 1\{x_{1,i0}\delta_1 + \alpha_{1,i} + \gamma_{1,0} + u_{1,i0} \geq 0\},\ \gamma_{1,0} \sim N(0,1/16), u_{1,i0} \sim N(0,1)$$

という形で決まるものとする（説明変数については $x_{j,i0} \sim N(0,1)$）。標本サイズや反復回数については表 1 の場合と同様に設定している。

表 2 については，（7）の係数 δ_2 と（9）の係数 ζ に結果に注目する。前者については，時間軸方

向への標本サイズが推定精度の大きな決定要因となっており，後者については時間軸方向への標本サイズと横断面方向のそれの比率も重要なバイアス発生要因となっているようである。つまり前者のδ_2については，時間軸方向の標本サイズが 50 を超えてくるような場合には無視できる程度のバイアスしか見られないが，後者のζ については，$N = T = 56$ の場合および $N = 100$, $T = 50$ の場合の比較からわかるように，横断面標本も時系列方向とともに大きくなる状況で初めてバイアスが消えている。

5. 結論

本稿では，パネルデータを用いた二段階推定量のバイアス修正について議論し，バイアス修正のための一般的な修正項の提示と標本選別モデルの場合における有限標本における振る舞いを見た。バイアス修正そのものは，修正前推定量を，偏り・効率性ともに明確に凌駕するが，いつでもバイアスを無視できるほど修正できるわけではなく，モデルや標本サイズに依存している部分もある。とくにそもそも推定量にバイアスが出やすい動学モデルのラグ付き従属変数の係数やコントロール関数の係数についてのバイアス修正は標本の性質に依存している部分も大きい。

本稿では，バイアス修正に関する一般的な公式を提示してはいるが，実際のデータ分析，とくに Heckman, Tobias, and Vytlacil［2003］のような処置効果推定におけるバイアス修正など，実証的に重要な問題については稿を改めて議論する。統計的な推測という観点からは標準誤差の頑健な推定も重要な課題である。Cameron and Miller［2015］では多重のクラスター効果を考慮した標準誤差の頑健推定の考察を行っており，複雑な固定効果を持つモデルへの応用へ示唆を与えている。また近年進展の著しいセミパラメトリックモデルの推定へパネルデータを利用する際（例えば，Li and Racine［2006］）のバイアス修正法へ拡張することは今後の課題として挙げることができる。

補論

（3）式における $\mathbf{b}_{1,i\bullet}$，$\mathbf{b}_{1,\bullet t}$ は以下のように与えられる。

$$\mathbf{b}_{1,i\bullet} = \mathbb{E}\left[\left(\sum_{t=1}^{T}\frac{\partial\tilde{\Psi}_{1,it}}{\partial\boldsymbol{\pi}_{it}}\right)\cdot\Delta_{1,i\bullet}\right] + \mathbb{E}\left[\sum_{t=1}^{T}\frac{\partial^2\tilde{\Psi}_{1,it}}{\partial\boldsymbol{\pi}_{it}\partial\boldsymbol{\pi}'_{it}}\right]\cdot\frac{\mathbb{E}[\Delta_{1,i\bullet}\otimes\Delta_{1,i\bullet}]}{2}$$

$$\mathbf{b}_{1,\bullet t} = \mathbb{E}\left[\left(\sum_{i=1}^{N}\frac{\partial\tilde{\Psi}_{1,it}}{\partial\boldsymbol{\pi}_{it}}\right)\cdot\Delta_{1,\bullet t}\right] + \mathbb{E}\left[\sum_{i=1}^{N}\frac{\partial^2\tilde{\Psi}_{1,it}}{\partial\boldsymbol{\pi}_{it}\partial\boldsymbol{\pi}'_{it}}\right]\cdot\frac{\mathbb{E}[\Delta_{1,\bullet t}\otimes\Delta_{1,\bullet t}]}{2}$$

また（4）式における $\mathbf{b}_{2,i\bullet}$，$\mathbf{b}_{2,\bullet t}$ は以下のように与えられる。

$$\begin{aligned}
\mathbf{b}_{2,i\bullet} &= \sum_{t=1}^{T}\mathbb{E}\left[\frac{\partial\Psi_{2,it}}{\partial\boldsymbol{\lambda}'_{it}}\frac{\partial\boldsymbol{\lambda}_{it}}{\partial\boldsymbol{\pi}'_{1,it}}\right]\cdot\boldsymbol{\beta}_{1,i\bullet} + \mathbb{E}\left[\left(\sum_{t=1}^{T}\frac{\partial\tilde{\Psi}_{2,it}}{\partial\boldsymbol{\pi}'_{2,it}}\right)\cdot\Delta_{2,i\bullet}\right] + \mathbb{E}\left[\left(\sum_{t=1}^{T}\frac{\partial\tilde{\Psi}_{2,it}}{\partial\boldsymbol{\lambda}'_{it}}\frac{\partial\boldsymbol{\lambda}_{it}}{\partial\boldsymbol{\pi}'_{1,it}}\right)\cdot\Delta_{1,i\bullet}\right] \\
&+\mathbb{E}\left[\left(\sum_{t=1}^{T}\frac{\partial\tilde{\Psi}_{2,it}}{\partial\boldsymbol{\pi}'_{1,it}}\right)\cdot\left(\sum_{t=1}^{T}\left(\sum_{s=1}^{T}\frac{\partial^2 q_{2,is}}{\partial\pi_{2,is}\partial\pi'_{2,is}}\right)^{-1}\frac{\partial^2 q_{2,it}}{\partial\pi_{2,it}\partial\lambda'_{it}}\frac{\partial\lambda_{it}}{\partial\pi_{1,it}}\right)\cdot\Delta_{1,i\bullet}\right] \\
&+\left(\sum_{t=1}^{T}\mathbb{E}\left[\frac{\partial^2\Psi_{2,it}}{(\partial\boldsymbol{\pi}_{2,it}\otimes\partial\boldsymbol{\pi}_{2,it})'}\right]\right)\cdot\frac{\mathbb{E}\left[\Delta_{2,i\bullet}\otimes\Delta_{2,i\bullet}\right]}{2} + \left(\sum_{t=1}^{T}\mathbb{E}\left[\frac{\partial^2\Psi_{2,it}}{(\partial\boldsymbol{\pi}_{2,it}\otimes\partial\boldsymbol{\lambda}_{it})'}\left(\mathbf{I}_{M_1}\otimes\frac{\partial\boldsymbol{\lambda}_{it}}{\partial\boldsymbol{\pi}'_{1,it}}\right)\right]\right)\cdot\mathbb{E}\left[\Delta_{2,i\bullet}\otimes\Delta_{1,i\bullet}\right] \\
&+\left(\sum_{t=1}^{T}\mathbb{E}\left[\frac{\partial\Psi_{2,it}}{\partial\boldsymbol{\lambda}'_{it}}\frac{\partial^2\boldsymbol{\lambda}_{it}}{(\partial\boldsymbol{\pi}_{1,it}\otimes\partial\boldsymbol{\pi}_{1,it})'} + \frac{\partial^2\Psi_{2,it}}{(\partial\boldsymbol{\lambda}_{it}\otimes\partial\boldsymbol{\lambda}_{it})'}\left(\frac{\partial\boldsymbol{\lambda}_{it}}{\partial\boldsymbol{\pi}'_{1,it}}\otimes\frac{\partial\boldsymbol{\lambda}_{it}}{\partial\boldsymbol{\pi}'_{1,it}}\right)\right]\right)\cdot\frac{\mathbb{E}\left[\Delta_{1,i\bullet}\otimes\Delta_{1,i\bullet}\right]}{2}
\end{aligned}$$

$$
\begin{aligned}
\mathbf{b}_{2,\bullet t} &= \sum_{i=1}^{N}\mathbb{E}\left[\frac{\partial\Psi_{2,it}}{\partial\boldsymbol{\lambda}'_{it}}\frac{\partial\boldsymbol{\lambda}_{it}}{\partial\boldsymbol{\pi}'_{1,it}}\right]\cdot\boldsymbol{\beta}_{1,\bullet t}+\mathbb{E}\left[\left(\sum_{i=1}^{N}\frac{\partial\tilde{\Psi}_{2,it}}{\partial\boldsymbol{\pi}'_{2,it}}\right)\cdot\Delta_{2,\bullet t}\right]+\mathbb{E}\left[\left(\sum_{i=1}^{N}\frac{\partial\tilde{\Psi}_{2,it}}{\partial\boldsymbol{\lambda}'_{it}}\frac{\partial\boldsymbol{\lambda}_{it}}{\partial\boldsymbol{\pi}'_{1,it}}\right)\cdot\Delta_{1,\bullet t}\right]\\
&+\mathbb{E}\left[\left(\sum_{i=1}^{N}\frac{\partial\tilde{\Psi}_{2,it}}{\partial\boldsymbol{\pi}'_{1,it}}\right)\cdot\left(\sum_{i=1}^{N}\left(\sum_{j=1}^{N}\frac{\partial^2 q_{2,jt}}{\partial\pi_{2,jt}\partial\pi'_{2,jt}}\right)^{-1}\frac{\partial^2 q_{2,it}}{\partial\pi_{2,it}\partial\lambda'_{it}}\frac{\partial\lambda_{it}}{\partial\pi_{1,it}}\right)\cdot\Delta_{1,\bullet t}\right]\\
&+\left(\sum_{i=1}^{N}\mathbb{E}\left[\frac{\partial^2\Psi_{2,it}}{(\partial\boldsymbol{\pi}_{2,it}\otimes\partial\boldsymbol{\pi}_{2,it})'}\right]\right)\cdot\frac{\mathbb{E}\left[\Delta_{2,\bullet t}\otimes\Delta_{2,\bullet t}\right]}{2}+\left(\sum_{i=1}^{N}\mathbb{E}\left[\frac{\partial^2\Psi_{2,it}}{(\partial\boldsymbol{\pi}_{2,it}\otimes\partial\boldsymbol{\lambda}_{it})'}\left(\mathbf{I}_{M_1}\otimes\frac{\partial\boldsymbol{\lambda}_{it}}{\partial\boldsymbol{\pi}'_{1,it}}\right)\right]\right)\cdot\mathbb{E}\left[\Delta_{2,\bullet t}\otimes\Delta_{1,\bullet t}\right]\\
&+\left(\sum_{i=1}^{N}\mathbb{E}\left[\frac{\partial\Psi_{2,it}}{\partial\boldsymbol{\lambda}'_{it}}\frac{\partial^2\boldsymbol{\lambda}_{it}}{(\partial\boldsymbol{\pi}_{1,it}\otimes\partial\boldsymbol{\pi}_{1,it})'}+\frac{\partial^2\Psi_{2,it}}{(\partial\boldsymbol{\lambda}_{it}\otimes\partial\boldsymbol{\lambda}_{it})'}\left(\frac{\partial\boldsymbol{\lambda}_{it}}{\partial\boldsymbol{\pi}'_{1,it}}\otimes\frac{\partial\boldsymbol{\lambda}_{it}}{\partial\boldsymbol{\pi}'_{1,it}}\right)\right]\right)\cdot\frac{\mathbb{E}\left[\Delta_{1,\bullet t}\otimes\Delta_{1,\bullet t}\right]}{2}
\end{aligned}
$$

最後に$\boldsymbol{\beta}_1$については，

$$
\begin{aligned}
\underbrace{\boldsymbol{\beta}_1}_{M_1(N+T)\times 1} &= \overline{\mathcal{H}}_1^{-1}\mathbb{E}\left[\widetilde{\mathcal{H}}_1\cdot\Delta_{1,NT}\right]+\overline{\mathcal{H}}_1^{-1}\mathbb{E}\left[\mathcal{Q}_{1,\phi\phi\phi}\right]\cdot\frac{\mathbb{E}\left[\Delta_{1,NT}\otimes\Delta_{1,NT}\right]}{2}\\
&\approx\begin{pmatrix}\left\{\left(\sum_{t=1}^{T}\left(\sum_{s=1}^{T}\frac{\partial q_{1,is}^2}{\partial\boldsymbol{\pi}_{1,is}\partial\boldsymbol{\pi}'_{1,is}}\right)^{-1}\frac{\partial q_{1,it}^2}{\partial\boldsymbol{\pi}_{1,it}\partial\boldsymbol{\pi}'_{1,it}}\right)\cdot\Delta_{1,i\bullet}+\sum_{t=1}^{T}\left(\sum_{s=1}^{T}\frac{\partial q_{1,is}^2}{\partial\boldsymbol{\pi}_{1,is}\partial\boldsymbol{\pi}'_{1,is}}\right)^{-1}\frac{\partial^2 q_{1,it}}{\partial\boldsymbol{\pi}_{1,it}(\partial\boldsymbol{\pi}_{1,it}^{\otimes 2})'}\frac{\mathbb{E}[\Delta_{1,i\bullet}\otimes\Delta_{1,i\bullet}]}{2}\right\}_{i=1}^{N}\\
\left\{\left(\sum_{i=1}^{N}\left(\sum_{j=1}^{N}\frac{\partial q_{1,jt}^2}{\partial\boldsymbol{\pi}_{1,jt}\partial\boldsymbol{\pi}'_{1,jt}}\right)^{-1}\frac{\partial q_{1,it}^2}{\partial\boldsymbol{\pi}_{1,it}\partial\boldsymbol{\pi}'_{1,it}}\right)\cdot\Delta_{1,\bullet t}+\sum_{i=1}^{N}\left(\sum_{j=1}^{N}\frac{\partial q_{1,jt}^2}{\partial\boldsymbol{\pi}_{1,jt}\partial\boldsymbol{\pi}'_{1,jt}}\right)^{-1}\frac{\partial^2 q_{1,it}}{\partial\boldsymbol{\pi}_{1,it}(\partial\boldsymbol{\pi}_{1,it}^{\otimes 2})'}\frac{\mathbb{E}[\Delta_{1,\bullet t}\otimes\Delta_{1,\bullet t}]}{2}\right\}_{t=1}^{T}\end{pmatrix}
\end{aligned}
$$

と書くことができる。ただし，$\partial\boldsymbol{\pi}_{1,it}^{\otimes 2}$は$\partial\boldsymbol{\pi}_{1\ it}\otimes\partial\boldsymbol{\pi}_{1,it}$を意味するものとする。また，最初の$N$要素が$\boldsymbol{\beta}_{1,i\bullet}$に対応し，残りの$T$要素が$\boldsymbol{\beta}_{1,\bullet t}$に対応している。

参考文献

Arellano, M., and J. Hahn［2006］“A likelihood-based approximate solution to the incidental parameter problem in dynamic nonlinear models with multiple effects” Available at http://www.cemfi.es/arellano/arellano-hahn-paper2006.pdf.

Bai, J.［2009］“Panel data models with interactive fixed effects”, *Econometrica*, Vol. 77, pp. 1229–1279.

Cameron, A.C., and D.L. Miller［2014］“Robust inference for dyadic data”, Available at http://cameron.econ.ucdavis.edu/research/dyadic_cameron_miller_december2014_with_tables.pdf

————, and ———［2015］“A practitioner's guide to cluster-robust inference”, *Journal Human Resources,* Vol. 50, pp. 317–372.

Chen, M.,I. Fernández-Val, and M. Weidner［2014］“Nonlinear panel models with interactive effects”, Available at arXiv:1412.5647.

Chudik, A., and H. Pesaran［2015］“Large panel data models with cross-sectional dependence: A survey”, in Baltagi, B.H. ed. The Oxford Handbook of Panel Data, Oxford: Oxford University Press, pp. 20–40.

Fernández-Val, I.［2009］“Fixed effects estimation of structural parameters and marginal effects in panel probit models”, Journal of Econometrics, Vol. 150, pp. 71–85.

————, and F. Vella［2011］“Bias corrections for two-step fixed effects panel data estimators”, Journal of

Econometrics, Vol. 163, pp. 144–162.

————————, and M. Weidner［2016］“Individual and time effects in nonlinear panel models with large N, T”, Journal of Econometrics, Vol. 192, pp. 291–312.

Hahn, J., and G. Kuersteiner［2011］“Bias reduction for dynamic nonlinear panel models with fixed effects”, Econometric Theory, Vol. 27, pp. 1152–1191.

————, and W. Newey［2004］“Jackknife and analytical bias reduction for nonlinear panel models”, Econometrica, Vol. 72, pp. 1295–1319.

Heckman, J., J.L. Tobias, and E. Vytlacil［2003］“Simple estimators for treatment parameters in a latent-variable framework”, Review of Economics and Statistics, Vol. 85, pp. 748–755.

————————［1976］“The common structure of statistical models of truncation, sample selection and limited Dependent Variables and a simple estimator for such models”, Annals of Economic and Social Measurement, Vol. 5, pp. 475–492.

Helpman, E., M. Melitz, and Y. Rubinstein［2008］“Estimating trade flows: Trading partners and trading volumes”, Quarterly Journal of Economics, Vol. 123, pp. 441–487.

Li, Q. and J.S. Racine［2006］Nonparametric Econometrics: Theory and Practice: Princeton University Press.

Maddala, G.S., and S. Wu［1999］“A comparative study of unit root tests with panel data and a new simple test”, Oxford Bulletin of Economics and Statistics, Vol. 61, pp. 631–665.

Newey, W., and K. West［1987］“A simple, positive semi-definite, heteroskedasticity and autocorrelation consistent covariance matrix”, Econometrica, Vol. 55, pp. 703–708.

Neyman, J., and E.L. Scott［1948］“Consistent estimates based on partially consistent observations”, Econometrica, Vol. 16, pp. 1–32.

Takara, Y.［2017］“Do Cultural Differences Affect the Trade of Cultural Goods? A Study in Trade of Music”, *Journal of Cultural Economics*,pp.1-25,2017,published online,DOI:10.1007/s 10824-017-9313-1.

Wooldridge, J.M.［2010］Econometric Analysis of Cross Section and Panel Data: The MIT Press.