6-2009

Discrete choice modeling with nonstationary panels applied to exchange rate regime choice

Sainan JIN
Singapore Management University, snjin@smu.edu.sg

Published in Journal of Econometrics, June 2009, Volume 150, Issue 2, pp. 312-321.
http://doi.org/10.1016/j.jeconom.2008.12.009

Discrete Choice Modeling with Nonstationary Panels Applied to Exchange Rate Regime Choice*

Sainan Jin
Guanghua School of Management
Peking University

August 23, 2007

Abstract

This paper develops a regression limit theory for discrete choice nonstationary panels with large cross section (N) and time series (T) dimensions. Some results emerging from this theory are directly applicable in the wider context of M-estimation. This includes an extension of work by Wooldridge (1994) on the limit theory of local extremum estimators to multi-indexed processes in nonlinear nonstationary panel data models.

It is shown that the ML estimator is consistent without an incidental parameters problem and has a limit theory with a fast rate of convergence $N^{1/2}T^{3/4}$ (in the stationary case, the rate is $N^{1/2}T^{1/2}$) for the regression coefficients and thresholds, and a normal limit distribution. In contrast, the limit distribution is known to be mixed normal in time series modeling, as shown in Park and Phillips (2000, hereafter PP), and Phillips, Jin, and Hu (2007, hereafter, PJH).

The approach is applied to exchange rate regime choice by monetary authorities, and we provide an analysis of the empirical phenomenon known as "fear of floating".
JEL Classification: C23, C25

Keywords: Discrete choice model, Exchange rate regime, Fear of floating, Fixed effects, Joint limits, Brownian local time.

*The author thanks Peter C. B. Phillips for his invaluable advice. Helpful comments and suggestions from the editors and the referees, Donald Andrews, Richard Baillie, Galina Hale, Jerry Hausman, Cheng Hsiao, Guido Kuersteiner, Kajal Lahiri, Vadim Marmer, Alex Maynard, Peter Robinson, Peter Schmidt, Kevin Song, Liangjun Su, Yixiao Sun, Yoon-Jae Whang, Halbert White, Jeffrey Wooldridge, the Yale econometrics reading group, and seminar participants at Michigan State University, SUNY-Albany, University of Chicago, UC Davis, UC Santa Cruz, University of Notre Dame, University of Pittsburgh, Yale University, 11th. International Panel Data Conference, 2005 International Symposium on Econometrics Development and 1st SETA conference are greatly appreciated. The author gratefully acknowledges the financial support from the NSFC (grants number: 70601001 and 70501001). Address Correspondence to: Sainan Jin, Guanghua School of Management, Peking University, Beijing, 100871, PR China. Email: sainan.jin@gsm.pku.edu.cn.

1 Introduction

Discrete choice panel data modeling has become a standard tool for empirical economic research. While traditional micro panel data empirical applications have been to large cross section (N) and fixed time series (T)[1], growing interest in cross-country analysis of macroeconomic policy decisions, currency crises and emerging stock market behavior has promoted the use of large dimensional panel data techniques. Often the time-series components exhibit strong evidence of nonstationarity. The goal of the present paper is to provide new asymptotics for such cases, in particular, for ordered discrete choice panel data regressions with individual effects that accommodate nonstationary data.

Phillips and Moon (1999, hereafter PM) developed a linear regression limit theory for nonstationary panels. Underlying their theory are asymptotics for multi-indexed processes in which both indexes may pass to infinity. This paper seeks to extend their limit theory to the maximum likelihood (ML) estimation of ordered discrete choice panel models. Since ML estimation in discrete choice models involves nonlinear optimization, we employ and further develop the asymptotic theory for nonlinear functions of integrated time series recently given in Park and Phillips (1999), PP, and PJH.

Panel models raise some additional problems that need to be addressed in this nonlinear nonstationary setting. Among these are the presence of an infinite dimensional space of fixed effects, the need for a multi-indexed $(N, T) \to \infty$ asymptotic theory of extremum estimation, and the possibility of multiple convergence rates. Some results emerging from this theory are directly applicable in the wider context of M-estimation. This includes an extension of work by Wooldridge (1994) on the limit theory of local extremum estimators to multi-indexed processes in nonlinear nonstationary panel data models.

In a panel discrete choice setting, we show that ML estimation provides consistent estimates of the full set of model parameters including the regression coefficients β_0, the thresholds μ_0 and the fixed effect α_{i0}. The ML estimator of the regression coefficients and thresholds has a normal limit distribution, whereas the limit distribution is known to be mixed normal in the time series case, as shown in PP and PJH. Another finding in this nonstationary setting is that fixed effects bias is removed asymptotically when $N/T^2 \to 0$, while in fixed effects nonlinear stationary panel modeling with large N large T, the fixed effects bias disappears asymptotically with the rate condition $N/T \to 0$, as shown in Hahn and Newey (2004).

The new results also stand in contrast to those of the binary choice models with zero thresholds where there are dual rates of convergence for the regression coefficients: a fast rate of convergence of $N^{1/2}T^{3/4}$ in a direction that is orthogonal to that of the true coefficient vector β_0; and a slower rate of convergence of $N^{1/2}T^{1/4}$ in other directions. This dual-rate phenomenon was discovered by PP in the time series binary choice setting. The difference in the asymptotic behavior in the present case

[1]See Baltagi (2005), Chamberlain (1984), Hsiao (2003), Matyas and Sevestre (1992) and Wooldrige (2002) for a review of traditional micro panel data literature.

arises from the fact that, for the ordered discrete choice case with nonstationary panels, we allow for scaled thresholds ($\sqrt{T}\mu$) consonant with the nonstationary nature of the data, and the signal from the regressors involves a nonlinear function of the covariates x_{it} evaluated at the linear form $x'_{it}\beta_0 - \sqrt{T}\mu_0^j$ instead of $x'_{it}\beta_0$. The latter (i.e. $x'_{it}\beta_0$) generally attenuates the signal from x_{it} in the direction β_0 because large deviations of $x'_{it}\beta_0$ enter as arguments of a density function which downweights large deviations, so they contribute less, as was pointed out in PP. On the other hand, the presence of scaled thresholds helps prevent the attenuation of the signal along β_0 because they recenter the main contribution to the signal at a spatial point away from the origin, thereby assuring the same rate of convergence in all directions. In addition, with scaled thresholds, the asymptotics involve functionals of local time at the thresholds instead of zero.

We apply our approach to model the choice of exchange rate regime by monetary authorities, and we explore the empirical phenomenon known as "fear of floating", which occurs when countries report a floating regime but actually intervene to smooth exchange rate fluctuations. This latter phenomenon was discussed in Calvo and Reinhart (2002) and has attracted much recent attention in international finance. We show that, consistent with the existing literature, fixed regimes are preferred by countries with smaller size, weaker government, more concentration in trade, and more foreign denominated liabilities. Also, countries that undergo a rapid process of financial deepening favor a more flexible exchange rate. We further show that fear of floating is positively associated with foreign denominated liabilities, and monetary shocks, among other variables.

The remainder of the paper is organized as follows. The next section outlines the discrete choice panel model with individual effects and assumptions. Section 3 gives the main results on the limit theory of the ML estimator. Section 4 presents an application to exchange rate regime choice. Section 5 concludes. Some useful lemmas and proofs of the main theorems are given in the Appendix of Jin (2006). Notation is given in a table at the end of the paper.

2 Basic Model, Assumptions

We will start with

$$y_{it}^* = \alpha_{i0} + x'_{it}\beta_0 - \varepsilon_{it}, \quad \text{for} \quad t = 1, ..., T, \text{ and } i = 1, ...N \tag{1}$$

where α_{i0} is the unobserved individual specific effect, x_{it} is an $m \times 1$ vector of explanatory variables and ε_{it} is an error. The dependent variable y_{it}^* is unobserved. Instead, what is observed is the indicator y_{it}, which takes the following possible $(J+1)$ values

$$
\begin{aligned}
y_{it} &= 0 \quad \text{if} \quad y_{it}^* \in (-\infty, \sqrt{T}\mu_0^1] \\
&= 1 \quad \text{if} \quad y_{it}^* \in (\sqrt{T}\mu_0^1, \sqrt{T}\mu_0^2] \\
&\vdots \\
&= J-1 \quad \text{if} \quad y_{it}^* \in (\sqrt{T}\mu_0^{J-1}, \sqrt{T}\mu_0^J] \\
&= J \quad \text{if} \quad y_{it}^* \in (\sqrt{T}\mu_0^J, \infty).
\end{aligned} \tag{2}
$$

Following PJH, the threshold parameters in (2) are scaled by $\sqrt{T}$ so that the thresholds have the same order of magnitude as the dependent variable y_{it}^* in (1) when the time series components of x_{it} are integrated processes. This avoids trivial asymptotic results and means, in effect, that the threshold levels adjust according to the sample size of the data. This seems realistic in a model where the covariates are allowed to be recurrent time series like integrated processes. We assume x_{it} is predetermined, i.e., $x_{i,t+1}$ is adapted to some filtration $(\mathcal{F}_t^i)$ with respect to which ε_{it} is measurable. In addition, following standard parametric discrete choice panel data modeling, we assume that ε_{it} is *i.i.d.* across i and t conditional on $(\mathcal{F}_{t-1}^i)$ with marginal distribution F, which is assumed to be known and standardized like a standard normal or the standard logistic. The model given by (1) and (2) is taken as correctly specified. The parameters of interest are assembled in the vector γ, whose true value $\gamma_0 = (\beta_0', \mu_0')'$ is an interior point of a subset of R^{m+J} which we assume to be compact and convex.

In the ordered discrete choice panel model with error distribution F, the conditional probability distribution of y_{it}, $P(y_{it} = j|\mathcal{F}_{t-1}^i) := P_j(x_{it}; \alpha_{i0}, \gamma_0)$ is given by

$$
\begin{aligned}
P_0(x_{it}; \alpha_{i0}, \gamma_0) &= 1 - F(\alpha_{i0} + x_{it}'\beta_0 - \sqrt{T}\mu_0^1), \\
P_J(x_{it}; \alpha_{i0}, \gamma_0) &= F(\alpha_{i0} + x_{it}'\beta_0 - \sqrt{T}\mu_0^J), \\
P_j(x_{it}; \alpha_{i0}, \gamma_0) &= F(\alpha_{i0} + x_{it}'\beta_0 - \sqrt{T}\mu_0^j) - F(\alpha_{i0} + x_{it}'\beta_0 - \sqrt{T}\mu_0^{j+1})
\end{aligned}
$$

for $j = 1, ..., J-1$.

The corresponding conditional expectation of y_{it} is

$$
\begin{aligned}
m(x_{it}; \alpha_{i0}, \gamma_0) &= \sum_{j=0}^{J} j \cdot P_j(x_{it}; \alpha_{i0}, \gamma_0) \\
&= \sum_{j=1}^{J} F(\alpha_{i0} + x_{it}'\beta_0 - \sqrt{T}\mu_0^j).
\end{aligned}
$$

If u_{it} is defined as the residual in the equation

$$
y_{it} = m_{it} + u_{it} = \sum_{j=1}^{J} F(\alpha_{i0} + x_{it}'\beta_0 - \sqrt{T}\mu_0^j) + u_{it}, \tag{3}
$$

then $(u_{it}, \mathcal{F}_t^i)$ is a martingale difference with conditional moments

$$\begin{aligned}\sigma_{k,it} &= \sigma_k(x_{it}; \alpha_{i0}, \gamma_0) \\ &= E(u_{it}^k | \mathcal{F}_{t-1}^i) \\ &= \sum_{j=0}^{J} (j - m_{it})^k \cdot P_j(x_{it}; \alpha_{i0}, \gamma_0), \text{ say.}\end{aligned}$$

Define $z_{k,it}$ as $z_k(x_{it}; \alpha_{i0}, \gamma_0) = u_{it}^k - \sigma_{k,it}$, $\eta_{kl,it}$ as $\eta_{kl}(x_{it}; \alpha_{i0}, \gamma_0) = E(z_{k,it} \cdot z_{l,it} | \mathcal{F}_{t-1}^i)$, and $a_{kl,it}$ as $a_{kl}(x_{it}; \alpha_{i0}, \gamma_0) = z_{k,it} z_{l,it} - \eta_{kl,it}$. Then $(z_{k,it}, \mathcal{F}_t^i)$, $(a_{kl,it}, \mathcal{F}_t^i)$ are also martingale differences. Obviously, $\sigma_{1,it} = 0$ and $z_{1,it} = u_{it}$. Further, define $\tau_{klpq,it} = E(a_{kl,it} \cdot a_{pq,it} | \mathcal{F}_{t-1}^i)$, giving fourth conditional moments for $z_{k,it.}$

For our asymptotic development we need more precise assumptions on the process generating x_{it}, and the following assumption is helpful.

Assumption 1: *Let $x_{it} = x_{it-1} + v_{it}$ with $x_{i0} = 0$ and where*

$$v_{it} = \sum_{s=0}^{\infty} C_{is} e_{it-s}$$

where

(a) *$\{C_{is}\}$ is a double sequence of $(m \times m)$ random matrices, i.i.d. across i for all s.*

(b) $\sum_{s=0}^{\infty} sE \parallel C_{is} \| < \infty$.

(c) *The innovations e_{it} are i.i.d. across i and over t with mean zero and $E(e_{it}e_{it}') = I_m$ and $E \parallel e_{it} \parallel^r < \infty$ for some $r > 8$, have a distribution that is absolutely continuous with respect to Lebesgue measure and have characteristic function φ_{it} which satisfies $lim_{\|t\| \to \infty} \|t\|^\kappa \varphi_{it} = 0$ for some $\kappa > 0$ and $i = 1, 2, ..., N$.*

(d) *C_{is} and e_{jt} are independent for all i, j, t and s.*

Remarks

(a) From Assumption 1, we have $C_i(1) = \sum_{s=0}^{\infty} C_{is} < \infty$ *a.s.* and *i.i.d.* across i. The assumption of *i.i.d.* random coefficient and the independence between C_{is} and e_{jt} are not necessary, we can relax this assumption easily with more complicated proofs.

(b) The linear process structure and the moment conditions on the innovations make available the use of embedding arguments and for which $\frac{1}{\sqrt{T}} x_{i[T.]} \Rightarrow V_i(\cdot) = C_i(1) W_i(\cdot)$ in $D[0,1]^m$, the m-fold Cartesian product of the space $D[0,1]$ endowed with the modified Skorohod metric (see Billingsley, 1999), where $C_i(1) = \sum_{s=0}^{\infty} C_{is} < \infty$ *a.s.* under Assumption 1, and V_i is a vector Brownian motion in $(\Omega^i, \mathcal{F}^i, P^i)$ with variance matrix $C_i(1)C_i(1)'$, and $W_i(\cdot)$ is a standard vector Brownian motion.

We also impose a mild restriction on the fixed effects parameters, which is useful in the proof of limit theory.

Assumption 2:

$$\sup_{1\leq i\leq N} |\alpha_{i0}| < M$$

for some $M < \infty$. As in PP, we rotate the regressor space to help isolate the effects of the nonlinearities. In particular, we assume that $\beta_0 \neq 0$ and rotate the regressor space using an orthogonal matrix $H = (h_1, H_2)$ with $h_1 = \beta_0/(\beta_0'\beta_0)^{1/2}$. Let $(\theta_0^1, \theta_0^{2\prime})' = \theta_0 = H'\beta_0$. Then we can write (1) as:

$$\begin{aligned} y_{it}^* &= \alpha_{i0} + x_{it}'\beta_0 - \varepsilon_{it} \\ &= \alpha_{i0} + x_{it}'HH'\beta_0 - \varepsilon_{it} \\ &= \alpha_{i0} + (H'x_{it})'H'\beta_0 - \varepsilon_{it} \\ &= \alpha_{i0} + x_{1,it}\theta_0^1 + x_{2,it}'\theta_0^2 - \varepsilon_{it}, \end{aligned}$$

where

$$\begin{aligned} x_{1,it} &= h_1'x_{it} \quad \text{and} \quad x_{2,it} = H_2'x_{it}, \\ \theta_0^1 &= h_1'\beta_0 = (\beta_0'\beta_0)^{1/2} \quad \text{and} \quad \theta_0^2 = H_2'\beta_0 = 0. \end{aligned}$$

Accordingly, we now define

$$V_{1i} = h_1'V_i \quad \text{and} \quad V_{2i} = H_2'V_i,$$

which are Brownian motions of dimensions 1 and $(m-1)$ respectively. Our subsequent theory involves the local time of the scalar process V_{1i}, which we denote by $L_{V_{1i}}(t,s)$, where t and s are the temporal and spatial parameters. $L_{V_{1i}}(t,s)$ is a stochastic process in time (t) and space (s) and represents the sojourn density of the process V_{1i} around the spatial point s over the time interval $[0,t]$. The reader is referred to Revuz and Yor (1994) for an introduction to the properties of local time and to Phillips (1998, 2001), PP for applications of this process in econometrics. In our analysis, it is more convenient to use the scaled local time of V_{1i} given by

$$L_{1i}(t,s) = (1/\sigma_{11,i})L_{V_{1i}}(t,s),$$

where $\sigma_{11,i}$ is the variance of V_{1i}.

Now we come back to the estimation of the ordered choice model. Let $\Lambda(i,t,j) = 1\{y_{it} = j\}$. The conditional log likelihood function can be written as:

$$\log L_{NT}(\alpha,\gamma) = \sum_{i=1}^{N} \log L_{iT}(\alpha_i,\gamma) = \sum_{i=1}^{N}\sum_{t=1}^{T}\sum_{j=0}^{J} \Lambda(i,t,j)\log P_j(x_{it};\alpha_i,\gamma).$$

Let the first derivative of F be denoted f and the second derivative be denoted $\dot{f}$. The scores are

$$\begin{aligned}
S_{\beta NT}(\alpha,\gamma) &= \sum_{i=1}^{N} S_{\beta iT}(\alpha_i,\gamma) = \frac{\partial \log L_{NT}}{\partial \beta} \\
&= \sum_{i=1}^{N}\sum_{t=1}^{T}\sum_{j=0}^{J} \frac{\Lambda(i,t,j)}{P_j(x_{it};\gamma)} p_j(x_{it};\alpha_i,\gamma) x_{it}, \\
S_{\mu^j NT}(\alpha,\gamma) &= \sum_{i=1}^{N} S_{\mu^j iT}(\alpha_i,\gamma) = \frac{\partial \log L_{NT}}{\partial \mu^j} \\
&= \sqrt{T}\sum_{i=1}^{N}\sum_{t=1}^{T}\left(\frac{\Lambda(i,t,j-1)}{P_{j-1}(x_{it};\alpha_i,\gamma)} - \frac{\Lambda(i,t,j)}{P_j(x_{it};\alpha_i,\gamma)}\right) f(\alpha_i + x_{it}'\beta - \sqrt{T}\mu^j), \\
S_{\alpha iT}(\alpha_i,\gamma) &= \frac{\partial \log L_{NT}}{\partial \alpha_i} = \sum_{t=1}^{T}\sum_{j=0}^{J} \frac{\Lambda(i,t,j)}{P_j(x_{it};\alpha_i,\gamma)} p_j(x_{it};\alpha_i,\gamma),
\end{aligned}$$

where

$$\begin{aligned}
p_0(x_{it};\alpha_i,\gamma) &= -f(\alpha_i + x_{it}'\beta - \sqrt{T}\mu^1), \\
p_J(x_{it};\alpha_i,\gamma) &= f(\alpha_i + x_{it}'\beta - \sqrt{T}\mu^J), \\
p_j(x_{it};\alpha_i,\gamma) &= f(\alpha_i + x_{it}'\beta - \sqrt{T}\mu^j) - f(\alpha_i + x_{it}'\beta - \sqrt{T}\mu^{j+1})
\end{aligned}$$

for $j = 1, \ldots, J-1$.

Note that the ratio $\frac{\Lambda(i,t,j)}{P_j(x_{it};\alpha_i,\gamma)}$ appears in all score functions. Since $E(\Lambda(i,t,j)|\mathcal{F}_{t-1}^i) = P_j(x_{it};\alpha_i,\gamma)$, the expected value of the ratio is 1. Let

$$\Lambda(i,t,j) = \frac{\prod_{s=0,\ldots,J \ \& \ s\neq j}(y_{it} - s)}{\prod_{s=0,\ldots,J \ \& \ s\neq j}(j - s)}, \tag{4}$$

then evaluated at true parameter values, the ratio can be written as a sum of martingale differences

$$\begin{aligned}
\frac{\Lambda(i,t,j)}{P_j(x_{it};\alpha_{i0},\gamma_0)} &= \frac{1}{P_j(x_{it};\alpha_{i0},\gamma_0)} \frac{\prod_{s=0,\ldots,J \ \& \ s\neq j}(y_{it} - s)}{\prod_{s=0,\ldots,J \ \& \ s\neq j}(j - s)} \\
&= \frac{1}{P_j(x_{it};\alpha_{i0},\gamma_0)} \frac{\prod_{s=0,\ldots,J \ \& \ s\neq j}(m_{it} + u_{it} - s)}{\prod_{s=0,\ldots,J \ \& \ s\neq j}(j - s)} \\
&= \sum_{k=1}^{J} g_k(x_{it};j,\alpha_{i0},\gamma_0)(u_{it}^k - \sigma_{k,it}) + 1 \\
&= \sum_{k=1}^{J} g_k(x_{it};j,\alpha_{i0},\gamma_0) z_{k,it} + 1,
\end{aligned}$$

where $g_k(x_{it};j,\alpha_{i0},\gamma_0)$ is defined to be the coefficient associated with $z_{k,it}$ for a given j and where $z_{k,it} = u_{it}^k - E(u_{it}^k|\mathcal{F}_{t-1}^i)$, which is a martingale difference. Using the above results, we have

$$
\begin{aligned}
S_{\beta iT}(\alpha_{i0},\gamma_0) &= \sum_{t=1}^{T}\sum_{k=1}^{J} A_{k,it} z_{k,it} x_{it}, && (5)\\
S_{\mu^j iT}(\alpha_{i0},\gamma_0) &= \sqrt{T}\sum_{t=1}^{T}\sum_{k=1}^{J} B_{k,it} z_{k,it}, && (6)\\
S_{\alpha iT}(\alpha_{i0},\gamma_0) &= \sum_{t=1}^{T}\sum_{k=1}^{J} A_{k,it} z_{k,it}, && (7)
\end{aligned}
$$

where

$$
\begin{aligned}
A_{k,it} &= \sum_{J=0}^{J} g_k(x_{it};j,\alpha_{i0},\gamma_0) p_j(x_{it};\alpha_{i0},\gamma_0)\\
&= \sum_{J=0}^{J} f(\alpha_{i0}+x_{it}'\beta_0-\sqrt{T}\mu_0^j)[g_k(x_{it};j,\alpha_{i0},\gamma_0)-g_k(x_{it};j-1,\alpha_{i0},\gamma_0)],\\
B_{k,it} &= f(\alpha_{i0}+x_{it}'\beta_0-\sqrt{T}\mu_0^j)[g_k(x_{it};j-1,\alpha_{i0},\gamma_0)-g_k(x_{it};j,\alpha_{i0},\gamma_0)].
\end{aligned}
$$

For the hessian, we have

$$
\begin{aligned}
S_{\beta\beta,NT}(\alpha,\gamma) &= \frac{\partial^2 \log L_{NT}}{\partial\beta\partial\beta'} = \sum_{i=1}^{N} S_{\beta\beta,iT}(\alpha_i,\gamma)\\
&= \sum_{i=1}^{N}\left(-\sum_{t=1}^{T}\sum_{k=1}^{J}\sum_{l=1}^{J} A_k A_l z_k z_l x_{it} x_{it}' + \sum_{t=1}^{T}\sum_{k=1}^{J} C_{\beta\beta,k} z_k x_{it} x_{it}'\right),
\end{aligned}
$$

$$
\begin{aligned}
S_{\beta\mu,NT}(\alpha,\gamma)(j) &= \frac{\partial^2 \log L_{NT}}{\partial\beta\partial\mu^j} = \sum_{i=1}^{N} S_{\beta\mu,iT}(\alpha_i,\gamma)(j)\\
&= \sum_{i=1}^{N}\left(-\sqrt{T}\sum_{t=1}^{T}\sum_{k=1}^{J}\sum_{l=1}^{J} A_k B_l(j) z_k z_l x_{it}' + \sqrt{T}\sum_{t=1}^{T}\sum_{k=1}^{J} C_{\beta\mu^j,k} z_k x_{it}'\right),
\end{aligned}
$$

$$
\begin{aligned}
S_{\beta\alpha,iT}(\alpha_i,\gamma) &= \frac{\partial^2 \log L_{NT}}{\partial\beta\partial\alpha_i}\\
&= -\sum_{t=1}^{T}\sum_{k=1}^{J}\sum_{l=1}^{J} A_k A_l z_k z_l x_{it}' + \sum_{t=1}^{T}\sum_{k=1}^{J} C_{\beta\alpha_i,k} z_k x_{it}',
\end{aligned}
$$

$$
\begin{aligned}
S_{\mu\mu,NT}(\alpha,\gamma)(j,j) &= \frac{\partial^2 \log L_{NT}}{\partial^2\mu^j} = \sum_{i=1}^{N} S_{\mu\mu,iT}(\alpha_i,\gamma)(j,j)\\
&= \sum_{i=1}^{N}\left(-T\sum_{t=1}^{T}\sum_{k=1}^{J}\sum_{l=1}^{J} B_k(j)B_l(j) z_k z_l - T\sum_{t=1}^{T}\sum_{k=1}^{J} C_{\mu^j\mu^j,k} z_k\right),
\end{aligned}
$$

$$\begin{aligned} S_{\mu\mu,NT}(\alpha,\gamma)(j,j-1) &= \frac{\partial^2 \log L_{NT}}{\partial\mu^j \partial\mu^{j-1}} = \sum_{i=1}^{N} S_{\mu\mu,iT}(\alpha_i,\gamma)(j,j-1) \\ &= \sum_{i=1}^{N} \left(-T \sum_{t=1}^{T} \sum_{k=1}^{J} \sum_{l=1}^{J} B_k(j) B_l(j-1) z_k z_l \right) \quad \text{for} \quad j = 2, ..., J, \end{aligned}$$

$$\begin{aligned} S_{\mu\mu,NT}(\alpha,\gamma)(j,j+1) &= \frac{\partial^2 \log L_{NT}}{\partial\mu^j \partial\mu^{j+1}} = \sum_{i=1}^{N} S_{\mu\mu,iT}(\alpha_i,\gamma)(j,j+1) \\ &= \sum_{i=1}^{N} \left(-T \sum_{t=1}^{T} \sum_{k=1}^{J} \sum_{l=1}^{J} B_k(j) B_l(j+1) z_k z_l \right) \quad \text{for} \quad j = 1, ..., J-1, \end{aligned}$$

$$\begin{aligned} S_{\mu\mu,NT}(\alpha,\gamma)(j,s) &= 0 \\ \text{for} \quad s &> j+1 \quad \text{and} \quad s < j-1, \end{aligned}$$

$$\begin{aligned} S_{\mu\alpha,iT}(\alpha_i,\gamma)(j) &= \frac{\partial^2 \log L_{NT}}{\partial\mu^j \partial\alpha_i} \\ &= -\sqrt{T} \sum_{t=1}^{T} \sum_{k=1}^{J} \sum_{l=1}^{J} B_k(j) A_l(i) z_k z_l + \sqrt{T} \sum_{t=1}^{T} \sum_{k=1}^{J} C_{\mu^j \alpha_i,k} z_k, \end{aligned}$$

$$S_{\alpha\alpha,iT}(\alpha_i,\gamma) = \frac{\partial^2 \log L_{NT}}{\partial^2 \alpha_i} = -\sum_{t=1}^{T} \sum_{k=1}^{J} \sum_{l=1}^{J} A_k A_l z_k z_l + \sum_{t=1}^{T} \sum_{k=1}^{J} C_{\alpha_i\alpha_i,k} z_k,$$

where we omit the arguments $(x_{it}; \alpha_i, \gamma)$ in the functions A, B, C and z for simplicity and where

$$\begin{aligned} C_{\beta\beta,k}(x_{it};\alpha_i,\gamma) &= C_{\beta\alpha_i,k}(x_{it};\alpha_i,\gamma) = C_{\alpha_i\alpha_i,k}(x_{it};\alpha_i,\gamma) \\ &= \sum_{j=0}^{J} \dot{p}_j(x_{it};\alpha_i,\gamma) g_k(x_{it};j,\alpha_i,\gamma), \\ C_{\beta\mu^j,k}(x_{it};\alpha_i,\gamma) &= C_{\alpha_i\mu^j,k}(x_{it};\alpha_i,\gamma) = \dot{p}_j(x_{it};\alpha_i,\gamma) g_k(x_{it};j,\alpha_i,\gamma), \\ C_{\mu^j\mu^j,k}(x_{it};\alpha_i,\gamma) &= \dot{f}(\alpha_i + x_{it}'\beta - \sqrt{T}\mu^j)(g_k(x_{it};j-1,\alpha_i,\gamma) - g_k(x_{it};j,\alpha_i,\gamma)). \end{aligned}$$

The ML estimator involves nonlinear functions of the integrated process x_{it} and it is helpful to be specific about the functions we need to consider. In the analysis below, we use the approach of PP and PJH in studying nonlinear transformations of integrated processes. A function f: $R \to R$ is called *regular* if it is bounded, integrable, and differentiable with bounded derivative. We denote by F_R the class of regular functions. We also consider the class F_I of bounded and integrable functions and the class F_0 of functions that are bounded and vanish at infinity. Clearly, $F_R \subset$

$F_I \subset F_0$. The following assumption about the distribution function F and density f of ε_{it} places uniform tail conditions on F and f, and places some explicit component functions in the classes. Both probit and logit functions satisfy conditions (a) - (c) and (8) of Assumption 3 (as discussed in PP and PJH), as is easily verified. The notation $\dot{g}$ and $\ddot{g}$ is used to denote the first and second derivatives of g.

Assumption 3: *F is three times differentiable with bounded derivatives and satisfies*

$$\sup_{|x|<M_m} \frac{F\left(x - M_m^{1+\eta}\mu\right)}{F(x)} = o(1), \quad \sup_{|x|\le M_m} \frac{1-F\left(x + M_m^{1+\eta}\mu\right)}{1-F(x)} = o(1),$$

$$\sup_{|x|<M_m} \frac{f\left(x \pm M_m^{1+\eta}\mu\right)}{f(x)} = o(1), \tag{8}$$

as $M_m \to \infty$ for any $\eta, \mu > 0$. *Further, for $k, l = 1, \ldots, J$:*

(a) $\eta_{kl}A_kB_l, \eta_{kl}A_kA_l, \eta_{kl}B_kB_l \in F_R$,

(b) $\eta_{kk}A_k, \eta_{kk}B_k, (\eta_{kl}\dot{A}_kB_l), (\eta_{kl}\dot{A}_kA_l), (\eta_{kl}\dot{B}_kB_l), \eta_{kk}^{1/2}\dot{C}_k \in F_I$,

(c) $\tau_{klpq}A_kA_lA_pA_q, \tau_{klpq}A_kA_lB_pB_q, \tau_{klpq}B_kB_lB_pB_q, \eta_{kl}C_kC_l \in F_0$.

3 Main Results

Corresponding to the rotation in the regressors and parameters, define

$$G = \begin{pmatrix} H & 0 \\ 0 & I_J \end{pmatrix}$$

and let $\rho = (\theta', \mu')'$. Then the score and hessian functions for the new parameter ρ are obtained from $S_{\rho NT}(\alpha, \rho) = G'S_{\gamma NT}(\alpha, \gamma)$ and $S_{\rho\rho,NT}(\alpha, \rho) = G'S_{\gamma\gamma,NT}(\alpha, \gamma)G$.

The estimator for the main parameter of interest ρ is obtained formally by concentrating out the fixed effects α_i from the log likelihood function. We solve

$$\widehat{\alpha}_i(\rho) \equiv \arg\max_{\alpha} \log L_{iT}(\alpha_i, \rho) \text{ and } \widehat{\rho} \equiv \arg\max_{\rho} \sum_{i=1}^{N} \log L_{iT}(\widehat{\alpha}_i(\rho), \rho)$$

where $\widehat{\alpha}_i(\rho)$ solves

$$S_{\alpha iT}(\widehat{\alpha}_i(\rho), \rho) = 0 \tag{9}$$

Using Lemmas 1, 2 and Theorem 1 in Jin (2006), we are able to characterize the asymptotic properties of the fixed effects estimator and the limit forms of the score and hessian functions of the regression coefficients and thresholds.

Theorem 1 *Let Assumptions 1-3 hold. Then as $(N, T) \to \infty$*

(a)

$$T^{-1/4}S_{\alpha iT}(\alpha_{i0}, \rho_0) \Rightarrow q_{44,i}^{1/2}W_i(1), \tag{10}$$

$$T^{-1/2}S_{\alpha\alpha,iT}(\alpha_{i0}, \rho_0) \Rightarrow -q_{44,i} \tag{11}$$

jointly, where $q_{44,i} = \sum_{j=1}^{J} \left\{ \frac{1}{\theta_0^1} L_{1i} \left(1, \frac{\mu_0^j}{\theta_0^1}\right) \int_{-\infty}^{\infty} \frac{f^2(s)}{F(s)(1-F(s))} ds \right\}$,

and W_i *is* $m-$*dimensional Brownian motion with covariance matrix* I, *which is independent of* V_i.

(b) *We could easily derive along the lines of the proof of Theorem R3 in PJH*

$$T^{1/4}(\widehat{\alpha}_i(\rho_0) - \alpha_{i0}) \Rightarrow q_{44,i}^{-1/2} W_i(1) \tag{12}$$

(c)

$$\begin{aligned}
N^{-1/2}T^{-3/4}S_{\rho NT}(\alpha_{i0},\rho_0) &= N^{-1/2}T^{-3/4}\sum_{i=1}^{N} S_{\rho iT}(\alpha_{i0},\rho_0) \Rightarrow N(0,Q_{11}), \\
N^{-1}T^{-3/2}S_{\rho\rho,NT}(\alpha_{i0},\rho_0) &= N^{-1}T^{-3/2}\sum_{i=1}^{N} S_{\rho\rho iT}(\alpha_{i0},\rho_0) \rightarrow_p -Q_{11}
\end{aligned}$$

jointly, where

$$Q_{11} = E(Q_{11,i}) = \begin{pmatrix} q_{11} & q_{12} & q_{13} \\ q_{21} & q_{22} & q_{23} \\ q_{31} & q_{32} & q_{33} \end{pmatrix}, Q_{11,i} = \begin{pmatrix} q_{11,i} & q_{12,i} & q_{13,i} \\ q_{21,i} & q_{22,i} & q_{23,i} \\ q_{31,i} & q_{32,i} & q_{33,i} \end{pmatrix} \tag{13}$$

with

$$\begin{aligned}
q_{11,i} &= \sum_{j=1}^{J} \left\{ \frac{(\mu_0^j)^2}{(\theta_0^1)^3} L_{1i}\left(1, \frac{\mu_0^j}{\theta_0^1}\right) \int_{-\infty}^{\infty} \frac{f^2(s)}{F(s)(1-F(s))} ds \right\}, \\
q_{12,i} &= \sum_{j=1}^{J} \left\{ \frac{\mu_0^j}{(\theta_0^1)^2} \int_0^1 dL_{1i}\left(r, \frac{\mu_0^j}{\theta_0^1}\right) V_{2i}(r)' \int_{-\infty}^{\infty} \frac{f^2(s)}{F(s)(1-F(s))} ds \right\}, \\
q_{13,i}(j) &= \frac{\mu_0^j}{(\theta_0^1)^2} L_{1i}\left(1, \frac{\mu_0^j}{\theta_0^1}\right) \int_{-\infty}^{\infty} \frac{f^2(s)}{F(s)(1-F(s))} ds, \\
q_{22,i} &= \sum_{j=1}^{J} \left\{ \frac{1}{\theta_0^1} \int_0^1 V_{2i}(r)V_{2i}(r)' dL_{1i}\left(r, \frac{\mu_0^j}{\theta_0^1}\right) \int_{-\infty}^{\infty} \frac{f^2(s)}{F(s)(1-F(s))} ds \right\}, \\
q_{23,i}(j) &= \frac{1}{\theta_0^1} \int_0^1 dL_{1i}\left(r, \frac{\mu_0^j}{\theta_0^1}\right) V_{2i}(r)' \int_{-\infty}^{\infty} \frac{f^2(s)}{F(s)(1-F(s))} ds, \\
q_{33,i}(j,j) &= \frac{1}{\theta_0^1} L_{1i}\left(1, \frac{\mu_0^j}{\theta_0^1}\right) \int_{-\infty}^{\infty} \frac{f^2(s)}{F(s)(1-F(s))} ds, \\
q_{33,i}(j,q) &= 0 \quad \text{for} \quad q \neq j.
\end{aligned}$$

To analyze the asymptotic behavior of parameters of interest, we follow Cox and Reid (1987) and Lancaster (2000, 2002) and reparameterize the model from $(\alpha', \rho')'$

to $\left(\lambda', \rho'\right)'$ so that ρ and λ_i are information orthogonal. Thus, $\alpha_i = \alpha(\lambda_i, \rho)$ is chosen such that the reparameterized log likelihood

$$\sum_{i=1}^{N} \log L_{iT}(\alpha_i, \rho) = \sum_{i=1}^{N} \log L_{iT}(\alpha(\lambda_i, \rho), \rho) = \sum_{i=1}^{N} \log L_{iT}^{*}(\lambda_i, \rho),$$

and

$$\log L_{iT}(\alpha_i, \rho) = \log L_{iT}^{*}(\lambda_i, \rho)$$

satisfies

$$E\left(\frac{\partial^2 \log^* L_{iT}}{\partial\rho\partial\lambda_i}\right) = 0$$

The orthogonalization treatment and asymptotic results for the score and hessian in Theorem 1 help deliver the limit distributions of $\widehat{\rho}$, which are established in Theorem 3.

General results for nonlinear nonstationary estimation in the case of panel data are not readily available in the literature. Theorem 10.1 of Wooldridge (1994) provided a standard approach to local extremum estimation for a single-indexed process. We extend this theorem for multi-indexed processes, which is suitable in nonlinear nonstationary panel data analysis.

Theorem 2 *(Joint Limits for Local Extremum Estimation in Panel: Extension of Wooldridge (1994, Theorem 10.1))*
Let $\{Q_{NT} : \mathcal{W} \times \Theta \to \mathbb{R}, N = 1, 2, ...; T = 1, 2, ...\}$ be a double-indexed sequence of objective functions defined on the data space $\mathcal{W}$ and the parameter space $\Theta \subset \mathbb{R}^P$ with score S_{NT} and hessian H_{NT}. Assume that

1. $\theta_0 \in int(\Theta)$;

2. *Q_{NT} satisfies the standard measurability and second order differentiability conditions on $\mathcal{W} \times \Theta$, $N = 1, 2, ...; T = 1, 2, ...$*

3. *There are sequences of nonstochastic positive definite diagonal matrices $\{\underline{C}_N\colon N = 1, 2, ...\}$, $\{C_T\colon T = 1, 2, ...\}$ and $\{\underline{D}_N\colon N = 1, 2, ...\}$ $\{D_T\colon T = 1, 2, ...\}$ such that the condition*

$$\max_{\mathcal{N}_{NT}^{o}} \| \underline{C}_N^{-1} C_T^{-1} [H_{NT}(\theta_0) - H_{NT}(\theta)] C_T^{-1} \underline{C}_N^{-1} \| = o_p(1) \tag{14}$$

holds with

$$\mathcal{N}_{NT}^{o} \equiv \{\theta \in \Theta : \| \underline{C}_N C_T (\theta - \theta_0) \| \leq 1\};$$

4. $\underline{C}_N C_T \underline{D}_N^{-1} D_T^{-1} \to 0$ *as* $(N, T) \to \infty$;

5. *After normalization, the score S_{NT} and hessian H_{NT} have joint limits:*

   (1) $\underline{D}_N^{-1} D_T^{-1} H_{NT}(\theta_0) D_T^{-1} \underline{D}_N^{-1} \to_p A_0$ as $(N, T) \to \infty$, where A_0 is a nonrandom, positive definite matrix;

(2) $\underline{D}_N^{-1} D_T^{-1} S_{NT}(\theta_0) \Rightarrow N(0, B_0)$ *as* $(N,T) \to \infty$, *where* B_0 *is a nonrandom, positive definite matrix.*

Then there exists a sequence of estimators $\left\{\widehat{\theta}_{NT} : N = 1,2,...; T = 1,2,...\right\}$ *such that*

$$S_{NT}(\widehat{\theta}_{NT}) = 0 \quad \text{w.p.a.l.} \tag{15}$$

$$\underline{D}_N D_T(\widehat{\theta}_{NT} - \theta_0) = -\left[\underline{D}_N^{-1} D_T^{-1} H_{NT}(\theta_0) D_T^{-1} \underline{D}_N^{-1}\right]^{-1} \underline{D}_N^{-1} D_T^{-1} S_{NT}(\theta_0) + o_p(1),$$

and

$$\underline{D}_N D_T(\widehat{\theta}_{NT} - \theta_0) \Rightarrow N(0, A_0^{-1} B_0 A_0^{-1}) \quad \text{as} \quad (N,T) \to \infty. \tag{16}$$

The above results help establish the form of the limit distribution of the ML estimator $\widehat{\rho}$, whose limit distribution is provided in the following.

Theorem 3 *Let Assumptions 1-3 hold. Then as* $N/T^{1/2} \to 0$

$$N^{1/2} T^{3/4} (\widehat{\rho} - \rho_0) \Rightarrow N\left(0, Q_{11}^{-1}\right),$$

where Q_{11} *is defined as in Theorem 1.*

Remarks

(a) As usual for local extremum estimation problems, Theorem 3 establishes the existence of a consistent root of the likelihood equation. The log likelihood function is well known to be concave in the logit and probit cases, and in such cases the consistent root is unique and is the global maximum.

(b) The ML estimator is consistent without an incidental parameters problem and has a normal limit distribution, while the limit distribution is mixed normal in time series modeling, as shown in PP and PJH. Further, estimators of the regression coefficients and thresholds converge at a faster rate $N^{1/2}T^{3/4}$ (in the stationary case, the rate is $N^{1/2}T^{1/2}$).

(c) The results also stand in contrast to those of the binary choice models with zero thresholds, where there are dual rates of convergence for the regression coefficients: a fast rate of convergence of $N^{1/2}T^{3/4}$ in a direction that is orthogonal to that of the true coefficient vector β_0; and a slower rate of convergence of $N^{1/2}T^{1/4}$ in other directions. This dual-rate phenomenon was discovered by PP in the time series binary choice setting. The difference in the asymptotic behavior in the present case arises from the fact that, for the ordered discrete choice case with nonstationary panels, we allow for scaled thresholds ($\sqrt{T}\mu$) consonant with the nonstationary nature of the data, and the signal from the regressors involves a nonlinear function of the covariates x_{it} evaluated at the linear form $x_{it}'\beta_0 - \sqrt{T}\mu_0^j$ instead of $x_{it}'\beta_0$. The latter (i.e. $x_{it}'\beta_0$) generally attenuates the signal from x_{it} in the direction β_0 because large deviations of $x_{it}'\beta_0$ enter as arguments of a density function which downweights large deviations, so they contribute less, as was pointed out in PP. On the other hand, the presence of scaled thresholds helps prevent the attenuation of the signal along β_0 because they recenter the main contribution to the signal at a spatial point away from the origin, thereby assuring the same rate of convergence in all directions. In addition, with scaled thresholds, the asymptotics involve functionals of local time at the thresholds instead of zero.

(d) In fixed effects nonlinear stationary panel modelling with large N large T, the fixed effects bias is removed when $N/T \to 0$, as shown in Hahn and Newey (2004). However, in this nonstationary setting, the rate condition becomes $N/T^{1/2} \to 0$. It is difficult to derive the explicit formula for the fixed effects bias or further conduct bias reduction when the time series component of x_{it} is an integrated process and the normed hessian converges weakly to a random limit matrix, not a constant matrix, as the time dimension passes to infinity.

(e) From Theorem 3, we have

$$N^{1/2}T^{3/4}G' \begin{pmatrix} \widehat{\beta} - \beta_0 \\ \widehat{\mu} - \mu_0 \end{pmatrix} \Rightarrow N(0, Q_{11}^{-1}),$$

that is,

$$N^{1/2}T^{3/4} \begin{pmatrix} \widehat{\beta} - \beta_0 \\ \widehat{\mu} - \mu_0 \end{pmatrix} \Rightarrow N(0, GQ_{11}^{-1}G'),$$

which we formalize as follows.

Corollary 1 *Under Assumptions 1-3, as* $N/T^{1/2} \to 0$

$$N^{1/2}T^{3/4} \begin{pmatrix} \widehat{\beta} - \beta_0 \\ \widehat{\mu} - \mu_0 \end{pmatrix} \Rightarrow N(0, GQ_{11}^{-1}G').$$

4 Application to Exchange Rate Regime Choice and Fear of Floating

Until recently, most of the empirical literature on exchange rate regimes utilized the IMF "*de jure*" classification published in IFS (International Financial Statistics), which required member states to self-declare their arrangements to IMF as belonging to one of four categories: peg, limited flexibility, managed floating, and independent floating. However, the classification of regimes is problematic. Deviations of actual behavior from announcements are very common. Recently, several attempts have been made to provide "*de facto*" classification. Levy-Yeyati and Sturzenegger (LS)

(2003) constructed a classification based on data on official exchange rates and international reserves using a cluster analysis methodology. Shambaugh (2004) offered another classification based on statistical analysis of the official exchange rate itself[2]. Reinhart and Rogoff (2004, hereafter RR) developed a novel system of re-classifying historical exchange rate regimes based upon a statistical analysis of the monthly data on market-determined parallel exchange rates and provide a fine classification with fourteen categories. They introduced a new category, free falling, for countries whose twelve-month rate of inflation is above forty percent. Traditionally free falling is labeled as independent floating, which is misleading, since the two regimes are different in important ways. They also offered a coarser classification that is conformable with that of IMF (except separating the free falling regimes from other floaters). Table 1[3] below displays a cross tabulation of observations in the IMF and RR classifications.

Table 1: The prevalence of deviations from announcements
Sample Period: 1974 – 2000

	Announcement (IMF)					
	Peg	Limited	Managed	Float	Inconcl.	Total
Actual (RR)						
Peg	803	94	80	33	5	1015
Limited	257	81	226	145	2	711
Managed	54	0	110	140	0	304
Float	12	25	9	125	1	172
Free fall	484	37	251	157	2	931
Inconcl.	0	0	0	1	0	1
Total	1610	237	676	601	10	3134

From Table 1, it is clear that there are significant differences between what countries say and what they do. A noticeable pattern is that there are deviations on both sides of the diagonals of the table, showing some countries peg more than they announced and others float more than they announced. Following Alesina and Wagner (2005), we define fear of floating as *de jure* floaters who intervene in the foreign exchange market to smooth the fluctuations of the nominal rate, as was also discussed in Calvo and Reinhart (2002), and fear of pegging as *de jure* fixers who are actually less fixed than they announced.

We begin with a brief analysis of the choice of exchange rate regimes using a multinomial ordered probit model[4]. Some recent literature argues that intermediate regimes for countries open to international capital flows are vanishing, the so-called "Bipolar view"[5]. We treat deviation regimes as intermediate regimes, and their

[2] Both LS and Shambaugh reclassifications are of annual frequency.

[3] Table 3 in Alesina and Wagner (2005).

[4] We could take the latent variable as the difference between costs and benefits of a fix. So the degree of the rigidity of the exchange rates will be continuous in that and so an ordered probit is appropriate.

[5] For recent discussions, please see Fischer (2001), Obstfeld and Rogoff (1995) and Summers (2000).

prevalence indicates that actually monetary authorities prefer an "interior solution". To put the point graphically, exchange rate arrangements lie along a line with hard pegs without deviation on the left and free and managed floating without deviation on the right. In the middle are intermediate regimes including soft pegs and deviation regimes. Thus, our dependent variable takes on three values, 0 for hard peg without deviation, 1 for intermediate regimes, and 2 for floating without deviation[6]. We use economic variables based on the exchange rate regime choice literature, especially Edwards (1996), Levy-Yeyati, Sturzenegger and Reggio (2003, hereafter LSR), and data availability. We use the lagged ratio of foreign denominated liabilities to money (FDLM) to control for the balance sheet effects. Country characteristics, such as openness (OPEN), size (SIZE) and geographical concentration of a country's trade (CTRADE), are included according to the traditional optimal currency area theory. To measure the importance of shocks we include terms of trade shocks (TTS), the volatility in the government consumption to GDP ratio (VGOV), and the volatility of money velocity (VVEL). We include two political variables to control for the strength of the government: the fraction of seats held in congress by the government party or coalition (MAJ), and the periods the incumbent administration has been in office (YRSOFF). The lagged ratio of quasi money over money (QMM) is selected as a proxy for the degree of domestic financial depth. The lagged ratio of international reserves relative to base money (RESBASE) is also included because a high level of reserves is often regarded as a necessary condition for the credibility and sustainability of a stable rate in developing countries. 50 countries in Africa, Asia, Europe, and Western Hemisphere constitute our sample[7]. Data sources and definitions are shown in Table 3.

We use quarterly data from 1975.1-2000.4[8]. Previous literature used unbalanced panels with annual frequency, which would suffer from an incidental parameters problem if we had fixed effects in the model. More importantly, the reclassification of exchange rate regimes by RR offered monthly data, and we find that each quarter IFS reports *de jure* exchange rate arrangement from 1974[9], thus it is more appropriate for us to use quarterly data[10]. The right two columns in Table 2 display the

[6] We also employ a different set of dependent variables, the value takes 0 for pegs (including soft pegs) without deviation, 1 for deviation regimes, and 2 for independent floating without deviation. The estimation results are similar.

[7] The countries are Argentina, Australia, Bolivia, Brazil, Burundi, Chile, Colombia, Costa Rica, Cyprus, Denmark, Dominica, Dominican Republic, Ecuador, Egypt, El Salvador, Guatemala, Haiti, Honduras, Iceland, India, Indonesia, Israel, Jamaica, Jordan, Kenya, Korea, Madagascar, Malawi, Malaysia, Mauritius, Mexico, Morocco, Nepal, New Zealand, Nicaragua, Nigeria, Pakistan, Paraguay, Peru, Philippines, Sri Lanka, Swaziland, Syrian Arab Republic, Tanzania, Thailand, Togo, Turkey, United states, Uruguay, and Venezuela.

[8] We exclude the monetary union countries in Europe (classified by IMF as limited flexibility) since the data for these countries are only available for the sample period 1974-1998. The estimation results are similar if we include the monetary union countries in our model.

[9] IFS offered quarterly exchange rate arrangements data starting from 1987. Before that, they published monthly arrangements.

[10] We thank Alesina and Wagner for sharing their annual dataset. We find that the estimation results are similar using their annual data with either LS or Shambaugh classification.

estimation results[11]. It is shown that the fixed rates are preferred by countries with smaller size, more concentration in trade, and more foreign denominated liabilities. Countries that undergo a rapid process of financial deepening favor a more flexible exchange rate, consistent with the argument of impossible trinity[12]. It seems that weak governments are prone to peg in order to gain credibility.

Next we provide an empirical analysis of fear of floating. It is not obvious why certain countries announce floating arrangements and then deviate. Recent literature, Reinhart, Rogoff and Savastano (2003), etc., has noted that liability dollarization, which is pervasive in emerging markets, may produce fear of floating. Hausman, Panizza, and Stein (2001) claimed that high pass-through and foreign currency liabilities tend to reduce the willingness of policymakers to let the exchange rate float freely. Calvo and Reinhart (2002) showed that fear of floating is pervasive, even among some of the developed countries. They pointed out that fear of floating arises from the combination of lack of credibility (as manifested in large and frequent risk-premia shocks), a high pass-through from exchange rates to prices, and inflation targeting. In Lahiri and Vegh (2001), output cost associated with exchange rate fluctuations gives rise to fear of floating. LSR argued that fear of floating appears to be associated with the prevalence of balance sheet effects and nominal shocks. Alesina and Wagner (2005) proposed a somewhat different hypothesis. In their view, devaluations may be perceived by the market as an indicator of turbulence and monetary instability. Thus, even countries that have claimed to be floaters may be induced to implement a peg in practice and fear of floating may be viewed as a signaling device to create confidence in the country.

We assume policy makers have an underlying utility function or optimal exchange rate regime choice rule that is not observed by econometricians

$$y_{it}^* = \alpha_{i0} + x_{it}'\beta_0 - \varepsilon_{it}, \text{ for } t = 1, ..., T, \text{ and } i = 1, ...N \tag{17}$$

Our dependent variable takes on two values: 1 when RR minus IMF is negative, i.e., when countries float less than what they say, and 0 otherwise. The choice of explanatory variables is driven by the analysis of the theoretical and empirical literature mentioned above, especially Calvo and Reinhart (2002), LSR, and by data availability. x_{it} is a vector of explanatory variables, which may be $I(0), I(d)$ or $I(1)$ processes or a mixture of these. Following LSR, we use the same exchange rate regime choice variables as above. Furthermore, lagged inflation (INF) is included as a proxy for the inflation targeting since monetary authorities may choose to peg in their attempts to lower inflation[13].

[11] Using *de jure or de facto* exchange rate arrangements yields similar estimation results.

[12] Impossible trinity has been discussed extensively in the international finance literature. Typically policy markers can choose at most two out of the three vortexes of the trinity: fixed exchange rates, independent monetary policy, or capital mobility.

[13] The link is controversial though, since persistent high inflation may create pressures on monetary authorities and force them to float.

We use quarterly data from 1975.1-2000.4. The first two columns in Table 2 show our estimation results for binary choice probit model[14] with and without the inclusion of fixed effects. In line with the existing literature, fear of floating is positively associated with foreign denominated liabilities, inflation, monetary shocks, but negatively correlated with international reserves. Openness, concentration in trade, terms of trade shocks, and quasi money ratio are not statistically significant, maybe because of the correlation between regressors. Political variables do not seem to play an important role, and with fixed effects the sign of YRSOFF changes.

Finally, we exclude free falling regimes from our sample, which are about 3 times as common as free floating regimes[15]. Given the distortions associated with very high inflation, any fixed versus flexible exchange rate regime comparisons that do not break out the free falling episodes may render meaningless results, as pointed out in RR. It turns out that the estimation results are similar[16], indicating the robustness of our findings.

5 Conclusion

Discrete choice models for panel data have proved to be a powerful tool in microeconometric analysis. The development of econometric theory that accommodates nonstationarity in nonlinear settings such as discrete choice opens up a range of potential applications in macroeconomics, finance, and international finance. The present paper seeks to provide such a theory for discrete choice panel regressions with individual effects that allow for nonstationary data. In particular, we provide a joint limit theory for the maximum likelihood estimation of such systems. The ML estimator is shown to be consistent and asymptotically normal without an incidental parameter problem and with faster rates of convergence for the regression coefficients and thresholds than in the stationary case.

Our approach is applied to modeling the choice of exchange rate regime by monetary authorities, and we provide an analysis of the phenomenon known as fear of floating. We show that countries with smaller size, weaker government, more concentration in trade, and more foreign denominated liabilities are more likely to peg. Also, countries that undergo a rapid process of financial deepening would like to adopt a more flexible exchange rate, consistent with the argument of impossible trinity. We further show that, consistent with the existing literature, fear of floating is associated with balance sheet effects, inflation targeting, and monetary shocks.

Hahn and Kuersteiner (2004) showed that as $N/T \to 0$ the fixed effects bias is removed asymptotically for dynamic nonlinear panel models with fixed effects. Of

[14] And again the estimation results are similar using annual data with either LS or Shambaugh classification.

[15] 19 countries remain in our sample. The countries are Australia, Colombia, Cyprus, Denmark, Dominica, Egypt, El Salvador, India, Malaysia, Mauritius, Morocco, Nepal, New Zealand, Pakistan, Sri Lanka, Swaziland, Syrian Arab Republic, Togo, and United states.

[16] To save space, we do not report the results here.

course, a different proof strategy is possible for nonstationary cases. Extensions along this line will be left for future research.

Table 2: Estimation Results
Sample Period: 1975.1 − 2000.4

	Fear of Floating		Regime Choice	
parameters	w/o fixed effects	fixed effects	w/o fixed effects	fixed effects
FDLM	3.6675* (0.243)	2.8537* (0.447)	−3.4211* (0.439)	−3.9077* (0.643)
OPEN	3.2336 (3.321)	3.7936 (3.279)	−1.9767 (1.429)	−2.3458 (1.927)
SIZE			7.0920* (2.364)	
CTRADE	−1.7901 (0.994)	−1.642 (1.091)	−1.7623* (0.209)	−1.0471* (0.236)
TTS	3.8796 (3.092)	4.6680 (3.293)	3.2674 (2.410)	4.2072 (3.614)
VGOV	4.7690* (1.357)	5.6242* (2.451)	−0.4671 (0.584)	−0.2312 (0.213)
VVEL	0.5476* (0.155)	0.6013* (0.200)	−0.0001 (0.001)	−0.0001 (0.001)
MAJ	−1.0662 (1.779)	−1.7846 (1.614)	1.4207* (0.433)	1.3096* (0.551)
YRSOFF	0.1221 (0.404)	−0.0856 (0.213)	0.0340* (0.011)	0.0731* (0.022)
QMM	0.1557 (0.204)	0.1876 (0.237)	0.5378* (0.089)	0.4480* (0.076)
RESBASE	−6.9746* (1.966)	−7.6529* (2.833)	−0.8766 (0.623)	−1.0413 (0.872)
INF	1.8671* (0.299)	1.9802* (0.477)	3.3457* (0.986)	3.1459* (1.452)
μ^1	0.1797* (0.036)	0.2458* (0.095)	0.0176* (0.002)	0.01904* (0.002)
μ^2			0.0442* (0.003)	0.0528* (0.003)

Standard errors in parentheses
$*$: significant at 5%

Table 3: Data Source (most definitions are from LSR (2002)

Variables	Definition and Source
FDLM	Lagged ratio of foreign liabilities to money (Source: IFS line 16C/line 34)
OPEN	Lagged openness (ratio of (exports+imports)/2 over GDP). (Source: IFS line 70/71, WDI)
SIZE	GDP in dollars over USA GDP (Source: WDI)
CTRADE	Lagged share of trade with the largest trading partner: exports to the largest trading partner as a share of total exports (Source: IMF-Direction of Trade Statistics)
TTS	Standard deviation of the logarithm of terms of trade over the previous 5 periods adjusted by openness (Source: WDI)
VGOV	Standard deviation of the government consumption to GDP ratio over the previous 5 periods (Source: IMF line 82/line 99b)
VVEL	Standard deviation of money velocity over the previous 5 periods (Source: IMF line 99b/line 34)
MAJ	Fraction of seats held by the government. (Source: Database of Political Institutions)
YRSOFF	Periods the incumbent administration has been in office (Source: Database of Political Institutions)
QMM	Lagged ratio of quasi money over money (Source: IMF line 35/line 34)
RESBASE	Lagged ratio of international reserves to monetary base (Source: IMF line 11/line 14)
INF	Lagged logarithm of one plus the percentage change in consumer price index (Source: IMF line 64)

6 Appendix: Useful Lemmas and Proofs

The corresponding appendix of Jin (2006) contains some useful lemmas, proofs of those lemmas, and proofs of the main results in the paper.

7 Notation

$\rightarrow_{a.s}$	almost sure convergence.
$\rightarrow_{p}$	convergence in probability.
$\Rightarrow, \rightarrow_{d}$	convergence in distribution.
$o_p(1)$	tends to zero in probability.
$=_d$	distributional equivalence.
$\sim_d$	asymptotically distributed as.
W, V_1, V_2	standard Brownian motions.
$L_V(t, s)$	local time of V at time t and spatial point s
$MN(0, V)$	mixed normal distribution with variance V.
$\Vert \cdot \Vert$	Euclidean norm in $\mathbf{R}^k$.
$\mathbf{F}_R$	class of regular functions.
$\mathbf{F}_I$	class of bounded integrable functions.
$\mathbf{F}_0$	class of bounded functions vanishing at infinity.

References

[1] Alesina, A. and A. Wagner (2005): "Choosing (and Reneging on) Exchange Rate Regimes," *Journal of the European Economic Association* 4, 770-799.

[2] Baltagi, B. (2005), *Econometric Analysis of Panel Data,* 3rd ed., John Wiley & Sons.

[3] Billingsley, P. (1999), *Convergence of Probability Measures*, 2nd ed., John Wiley & Sons.

[4] Calvo, G and C. Reinhart (2002): "Fear of Floating," *The Quarterly Journal of Economics,* 117, 379-408.

[5] Chamberlain, G. (1984): "Panel Data," in Z.Griliches and Intriligator (eds.), *Handbook of Econometrics,* vol. 2, Amsterdam, North-Holland.

[6] Cox, D. R. and N. Reid (1987): "Parameter Orthogonality and Approximate Conditional Inference," *Journal of the Royal Statistical Society*, Series B 49, 1-39.

[7] Edwards, S. (1996): "The Determinants of the Choice between Fixed and Flexible Exchange-Rate Regimes," NBER working paper 5756.

[8] Hahn, J. and W. Newey (2004): "Jackknife and Analytical Bias Reduction for Nonlinear Panel Models," *Econometrica*, 72, 1295-1320.

[9] Hahn, J. and G. Kuersteiner (2004): "Bias Reduction for Dynamic Nonlinear Panel Models with Fixed Effects," unpublished manuscript.

[10] Hausmann, R., U. Panizza, and E. Stein (2001): "Why Do Countries Float the Way They Float," *Journal of Development Economics*, Elsevier, vol. 66, 387-414.

[11] Hsiao, C. (2003), *Analysis of Panel Data,* 2nd ed., Cambridge University Press.

[12] Jin, S. (2006): "Discrete Choice Modeling with Nonstationary Panels Applied to Exchange Rate Regime Choice," working paper, Peking University.

[13] Lahiri, A. and C. Vegh (2001): "Living with the fear of floating: An Optimal Policy Perspective," in S. Edwards and J. Frankel (eds) Preventing Currency Crises in Emerging Markets, U of Chicago Press.

[14] Lancaster, T. (2000): "The Incidental Parameter Problem since 1948," *Journal of Econometrics* 95, 391-413.

[15] Lancaster, T. (2002): "Orthogonal Parameters and Panel Data," *Review of Economic Studies* 69, 647-666.

[16] Levy-Yeyati, E., F. Sturzenegger and I. Reggio (2003): "On the Endogeneity of Exchange Rate Regimes," unpublished manuscript.

[17] Levy-Yeyati, E., F. Sturzenegger (2003): "A de facto Classification of Exchange Rate Regimes: A Methodological Note," *American Economic Review* 93, 1173-1193.

[18] Matyas, L. and P. Sevestre (eds.) (1992), *The Econometrics of Panel Data*, Boston, MA: Kluwer Academic Publisher.

[19] Park, J. Y. and P. C. B. Phillips (1999): "Asymptotics for Nonlinear Transformations of Integrated Time Series," *Econometric Theory,* 15, 269-298.

[20] Park, J. Y. and P. C. B. Phillips (2000): "Nonstationary Binary Choice," *Econometrica,* 68, 1249-1280

[21] Phillips, P. C. B. (1998): "Econometric Analysis of Fisher's Equation," CFDP 1180, Yale University.

[22] Phillips, P. C. B. (2001): "Descriptive Econometrics for Non-Stationary Time Series with Empirical Applications," *Journal of Applied Econometrics,* 16, 389-413.

[23] Phillips, P. C. B., S. Jin and L. Hu (2007): "Nonstationary Discrete Choice: A Corrigendum and Addendum," forthcoming, *Journal of Econometrics..*

[24] Phillips, P. C. B. and H. R. Moon (1999): "Linear Regression Limit Theory for Nonstationary Panel Data," *Econometrica,* 67, 1057-1111.

[25] Reinhart, C. and K. Rogoff (2004): "The Modern History of Exchange Rate Arrangements: A Reinterpretation," *Quarterly Journal of Economics*, vol. 119, 1-48.

[26] Reinhart, C., K. Rogoff and M. Savastano (2003): "Addicted to Dollars," NBER Working Paper 10015.

[27] Revuz, D. and M.Yor (1994), *Continuous Martingale and Brownian Motion*, 2nd ed. Springer-Verlag: New York.

[28] Shambaugh, J (2004): "The Effect of Fixed Exchange Rates on Monetary Policy," *Quarterly Journal of Economics*, vol. 119, 301-352.

[29] Summers, L. H. (2000): "International Financial Crises: Causes, Prevention, and Cures," *American Economic Review, Papers and Proceedings,* 90, 1-16.

[30] Wooldridge, J. M. (1994): "Estimation and Inference for Dependent Processes," In R.F. Engle and D.L. McFadden (Eds.) *Handbook of Econometrics*, vol. 4, Amsterdam, North-Holland.

[31] Wooldridge, J. M. (2002), *Econometric Analysis of Cross Section and Panel Data,* MIT Press.