Stationarity and Mixing Properties of the Dynamic Tobit Model

Jinyong Hahn[a] Guido Kuersteiner[b]
UCLA UC Davis

Abstract

We establish strict stationarity and strong mixing properties of the dynamic Tobit process. Using these results we show that the regularity conditions for bias corrections in general non-linear dynamic panel models are satisfied for the dynamic Tobit model.

JEL: C13, C24

Keywords: Dynamic Tobit Model, stationarity, mixing.

---

[a] UCLA, Department of Economics, 8283 Bunche Hall, Mail Stop 147703, Los Angeles, CA 90095; (310) 825 2523; email: hahn@econ.ucla.edu

[b] Corresponding Author; Contact Information: UC Davis, Dept. of Economics, One Shields Ave, Davis, CA 95616; (916) 524 3428; email:gkuerste@ucdavis.edu

1 Introduction

The purpose of this paper is to establish strict stationarity and mixing properties for the dynamic Tobit process of the form

$$(1) \qquad y_{it} = \max\left(0, z_{it}'\zeta_0 + \tau_0 y_{it-1} + \gamma_{i0} + \varepsilon_{it}\right)$$

where z_{it} is an exogenous regressor, ε_{it} is an independent disturbance term and γ_{i0}, τ_0 and ζ_0 are fixed parameters. Tobin (1958) proposed a static version of Model (1) to analyze durable good expenditures. The model has subsequently found wide application in economics. However, most of the literature worked with a static version of the model and cross-sectional data or with panel data when the number of time series observations is assumed fixed. More recently some authors have considered time series versions of the models. Examples include de Jong and Herrera (2005) and Hahn and Kuersteiner (2004).

Stationarity and Mixing properties for linear processes are well understood and documented in an extensive literature. The same is true only to a much lesser extent for non-linear models. De Jong and Woutersen (2005) establish strict stationarity and mixing for certain discrete choice models but to the best of our knowledge mixing has not been established for (1). Because a mixing sequence retains the mixing property under any measurable transformation establishing mixing for a particular stochastic process is of independent interest. As our application of our results to the maximum likelihood estimator of (1) shows the main application of mixing properties in statistics are moment bounds to control the temporal dependence of expressions entering the likelihood.

The main purpose of this paper is to establish that the regularity conditions of Hahn and Kuersteiner (2004) are satisfied. In that paper we develop bias correction methods for non-linear dynamic panel models with fixed effects under a mixing condition. DeJong and Herrera (2005) analyze a time series version of the dynamic Tobit model and use the concept of near epoch dependence to derive asymptotic properties. While near epoch dependence is easier to establish than mixing, it does not satisfy the regularity conditions in Hahn and Kuersteiner (2004).

2 Stationarity and Mixing

The first result that we establish is that for innovation sequences ε_{it} that are independent both in the cross-section i and time series t and for inputs z_{it} which are independent of ε_{js} for all

$i \neq j$ and $t \neq s$ it follows that there exists a stationary solution y_{it} to equation (1). Our result applies both to the pure time series case where there is only one single cross-sectional unity such that the index i can be omitted or for the more general panel case where the time series is observed for multiple units in the cross-section. The result is formally stated as follows:

Proposition 1 *Let $i \in \mathbb{N}_+$ be an index of a countable, not necessarily finite number of cross-sectional units. Let ε_{it} be iid across i and t with $E\left|\varepsilon_{it}\right|^r < \infty$ for some $r \geq 2$ and z_{it} is strictly stationary and $\sup_i E\left|z_{it}\right|^r < \infty$. Then $y_{it} = \max(0, \tau_0 y_{it-1} + z_{it}'\zeta_0 + \gamma_{i0} + \varepsilon_{it})$ with $|\tau_0| < 1$ has a strictly stationary solution y_{it} and $\sup_i E\left|y_{it}\right|^r < \infty$.*

We now turn to the main result of our paper which is concerned with the mixing properties of stationary solutions to equation (1). We first define the mixing property in the context of our model. Let $x_{it} = (y_{it}, z_{it})$ where y_{it} is the stationary solution to (1). Let $\mathcal{A}_{it} = \sigma\left(x_{it}, x_{it-1}, x_{it-2}, ...\right)$ and $\mathcal{B}_{it} = \sigma\left(x_{it}, x_{it+1}, x_{it+2}, ...\right)$ be the sigma algebras generated by $\left(x_{it}, x_{it-1}, x_{it-2}, ...\right)$ and $\left(x_{it}, x_{it+1}, x_{it+2}, ...\right)$ respectively. Define

$$\alpha_i\left(m\right) = \sup_t \sup_{A \in \mathcal{A}_{it}, B \in \mathcal{B}_{it+m}} \left|P(A \cap B) - P\left(A\right) P\left(B\right)\right|$$

The process x_{it} is said to be strong mixing if $\sup_i \alpha_i\left(m\right)$ tends to 0 as $m \to \infty$. We establish not only that x_{it} is strong mixing but also that the mixing coefficient $\alpha_i\left(m\right)$ decays exponentially fast as $m \to \infty$. Our mixing result can now be stated as follows:

Proposition 2 *Let ε_{it} be iid across i and t with density $p_\varepsilon(\varepsilon)$ such that ε_{it} has unbounded support with $0 < P(\varepsilon_{it} < x) < 1$ for all $x \in \mathbb{R}$ and $\int \left|p_\varepsilon\left(\varepsilon - x\right) - p_\varepsilon(\varepsilon)\right| d\varepsilon \leq C_\varepsilon \left|x\right|$ and $E\left[\left|\varepsilon_{it}\right|^r\right] < \infty$ for some $r \geq 2$. Assume that $z_{it} = \Pi z_{it-1} + \eta_{it}$ where η_{it} is independent of ε_{js} for all i, j, t, s and that η_{it} has density $p_\eta(\eta)$ such that $\int \left|p_\eta\left(\eta - x\right) - p_\eta(\eta)\right| d\eta \leq C_\eta \left\|x\right\|$ and $E\left\|\eta_{it}\right\|^r < \infty$ for some $r > 2$. Let π_j be the eigenvalues of Π with $\max_j \left|\pi_j\right| < 1$. Then $y_{it} = \max(0, \tau_0 y_{it-1} + z_{it}'\zeta_0 + \gamma_{i0} + \varepsilon_{it})$ with $|\tau_0| < 1$ has a stationary solution y_{it}, $\sup_i E\left|y_{it}\right|^r < \infty$, $x_{it} = (y_{it}, z_{it})$ is mixing with $\sup_i \left|\alpha_i\left(m\right)\right| \leq C a^m$ for some a such that $|a| < 1$.*

Remark 1 *The smoothness conditions for $p_\varepsilon\left(\varepsilon\right)$ and $p_\eta\left(\eta\right)$ can not be easily relaxed. To show this consider a simple counter example. Let e_t be iid with distribution $P(e_t = 0) = P(e_t = 1) = 1/2$ and define the process $Y_t^* = 2^{-1} Y_{t-1}^* + e_t$ with initial condition $Y_0^* = 0$. It follows immediately that $Y_t^* \geq 0$ with probability one for all t. Then consider $Y_t = \max(0, 2^{-1} Y_{t-1} + e_t)$ with $Y_0 = 0$. Obviously, $Y_t = Y_t^*$ for all t and all realizations of e_t. But from Athreya and Pantula (1986) it is known that Y_t^* is not strong mixing and thus Y_t is not strong mixing. This is despite the fact that $|\tau_0| = 1/2 < 1$ in this example such that the mapping $\max(0, .)$ is a contraction.*

3 Maximum Likelihood Estimation

If we are prepared to impose parametric assumptions on the distribution of ε_{it} then the parameters of (1) can be estimated by the method of maximum likelihood for an observed sample $\{x_{it}\}_{i=1,t=1}^{N,T}$ generated from (1). For expositional purposes we assume that ε_{it} is iid Gaussian. In this case the log of the likelihood is given by

$$\psi\left(x_{it};\theta,\gamma_i\right) = 1\left(y_{it}=0\right)\log\Lambda\left(\left(\tau y_{it-1}+z_{it}'\zeta+\gamma_i\right)/\sigma_\varepsilon^2\right)+1\left(y_{it}>0\right)\lambda\left(\left(y_{it}-\tau y_{it-1}-z_{it}'\zeta-\gamma_i\right)/\sigma_\varepsilon^2\right),$$

where Λ is the cumulative distribution function of the standard normal distribution and λ is the corresponding density. It is well known that the maximum likelihood estimator obtained from maximizing $\sum_{i=1}^N\sum_{t=1}^T\psi\left(x_{it};\theta,\gamma_i\right)$ over $\{\theta,\gamma_1,...,\gamma_N\}$ is inconsistent if T is held fixed as $N\to\infty$. This problem is known as the incidental parameter problem and discovered by Neyman and Scott (1948). In Hahn and Kuersteiner (2004) we show under high level regularity conditions imposed on $\psi(.)$ and x_{it} that by adopting a joint limiting argument where $N,T\to\infty$ jointly such that $T/N\to\kappa$ where κ is a constant such that $0<\kappa<\infty$ one can establish the consistency of the maximum likelihood estimator $\hat{\theta}$. Moreover, we show that $\hat{\theta}$ is $\sqrt{TN}$ convergent and asymptotically normal but with a limiting distribution that is not centered at zero. We interpret the location shift of the limiting distribution as an approximation to the finite sample bias and show that successful bias correction is possible even in samples where T is substantially smaller than N.

Our bias approximation and correction results depend on high level assumptions, in particular, the existence of higher order moments and mixing and stationarity properties of x_{it}. The following result establishes that these conditions are satisfied for data generated by (1) under certain additional regularity conditions which are stated as part of the proposition.

Proposition 3 *Let ε_{it} be iid across i and t with $\varepsilon_{it}\sim N(0,\sigma_\varepsilon^2)$. Assume that $z_{it}=\Pi z_{it-1}+\eta_{it}$ where η_{it} is independent of ε_{js} for all i,j,t,s and that η_{it} has density $p_\eta(\eta)$ such that $\int\left|p_\eta\left(\eta-x\right)-p_\eta(\eta)\right|d\eta\leq C_\eta\left\|x\right\|$ and $E\left\|\eta_{it}\right\|^r<\infty$ for some $r>7+10q+12+\delta$ with $q\geq p/2+2$ and for some $\delta>0$. Let π_j be the eigenvalues of Π with $\max_j|\pi_j|<1$. Let $y_{it}=\max(0,\tau_0y_{it-1}+z_{it}'\zeta_0+\gamma_{i0}+\varepsilon_{it})$ and define $x_{it}=(y_{it},y_{it-1},z_{it})$ and*

$$\begin{aligned}\psi\left(x_{it};\theta,\gamma_i\right) &= 1\left\{y_{it}=0\right\}\log\Lambda\left(\left(\tau y_{it-1}+z_{it}'\zeta+\gamma_i\right)/\sigma_\varepsilon^2\right)\\ &\quad +1\left\{y_{it}>0\right\}\log\lambda\left(\left(y_{it}-\tau y_{it-1}-z_{it}'\zeta-\gamma_i\right)/\sigma_\varepsilon^2\right).\end{aligned}$$

Then there exists a function $M\left(x_{it}\right)$ such that

$$\left|D^v\psi\left(x_{it},\beta_1\right)-D^v\psi\left(x_{it},\beta_2\right)\right|\leq M\left(x_{it}\right)\left\|\beta_1-\beta_2\right\|$$

for all $\beta_1, \beta_2 \in \Phi$ *and* $|v| \leq 5$, $\sup_{\beta \in \Phi} \left\| D^v \psi\left(x_{it}, \beta\right) \right\| \leq M(x_{it})$ *and*

$$\sup_i E\left[\left|M(x_{it})\right|^{10q+12+\delta}\right] < \infty$$

for some integer $q \geq p/2 + 2$ *and for some* $\delta > 0$.

Let

$$\left(\widehat{\theta}, \widehat{\gamma}_1, \ldots, \widehat{\gamma}_n\right) = \underset{\theta, \gamma_1, \ldots, \gamma_n}{\text{argmax}} \sum_{i=1}^{N} \sum_{t=1}^{T} \psi\left(x_{it}; \theta, \gamma_i\right)$$

be the maximum likelihood estimator. It is easy to verify that the regularity conditions of Hahn and Kuersteiner (2004) are satisfied when the conclusions of Proposition 3 hold. The results of that paper then show that when $T, N \to \infty$ and $N/T \to \kappa$ the parameters are uniformly consistent. In other words, $\hat{\theta} - \theta_0 = o_p(1)$ and $\max_i \left(\hat{\gamma}_i - \gamma_{i0}\right) = o_p(1)$. In addition, it is shown in Hahn and Kuersteiner (2004) that a bias correction for the ML estimator can be based on an asymptotic approximation where N and T tend to infinity at the same rate.

4 Conclusions

We establish that the dynamic Tobit model satisfies stationarity and mixing properties under some regularity conditions. As an application of our results we show that the high level regularity assumptions of Hahn and Kuersteiner (2004) are satisfied.

Appendix

Proof of Proposition 1:. We note that $z^*_{it} = z'_{it}\zeta_0 + \gamma_{i0} + \varepsilon_{it}$ is strictly stationary. To show stationarity of y_{it} we adapt the argument of de Jong and Woutersen (2003). Let y^k_{it} be a process that satisfies $y^k_{is} = \max(0, \tau_0 y^k_{is-1} + z^*_{is})$ for $s = t, t-1, ..., t-k$ and $y^k_{is} = 0$ for $s < t-k$. Note that $y^k_{it} = f(z^*_{it}, ..., z^*_{it-k})$ for some measurable map f such that y^k_{it} is strictly stationary for all k by construction. By Theorem 6.10 of Pötscher and Prucha (1991) it follows that $\sup_i \sup_{t\geq 0} E\left[\left|y^k_{it}\right|^r\right] < \infty$ uniformly in k. It follows that $E\left[\left|y^k_{it} - y^{k+m}_{it}\right|^2\right] \leq \left|\tau_0\right|^{2k} E\left[\left|y^m_{it-k}\right|^2\right] \to 0$ as $k, m \to \infty$. This implies that y^k_{it} is a Cauchy sequence with a mean square limit denoted by y_{it}. By Fatou's Lemma, $E\left|y_{it} - \lim_{k\to\infty} y^k_{it}\right|^2 \leq \liminf_{k\to\infty} E\left|y_{it} - y^k_{it}\right|^2 \to 0$ such that y_{it} and $\lim_k y^k_{it}$ are equal almost surely. This implies that y_{it} is stationary.

Assume that y^*_{it} is a solution to $y_{it} = \max(0, \tau_0 y_{it-1} + z^*_{it})$. Then, $\left|y^*_{it}\right| \leq \sum_{j=0}^{\infty} \left|\tau_0\right|^j \left|z^*_{it-j}\right|$ which implies that $\sup_i \sup_{t\geq 0} E\left[\left|y^*_{it}\right|^r\right] < \infty$. Consider

$$\begin{aligned} \left|y^*_{it} - y^k_{it}\right| &= \left|\max(0, \tau_0 y^*_{it-1} + z^*_{it}) - \max(0, \tau_0 y^k_{it-1} + z^*_{it})\right| \\ &\leq \left|\tau_0\right| \left|y^*_{it-1} - y^k_{it-1}\right| \leq .... \leq \left|\tau_0\right|^k \left|y^*_{it-k}\right| \end{aligned} \tag{2}$$

and

$$\left|y^*_{it} - y_{it}\right| \leq \left|y^*_{it} - y^k_{it}\right| + \left|y_{it} - y^k_{it}\right| \leq \left|\tau_0\right|^k \left|y^*_{it-k}\right| + \left|y_{it} - y^k_{it}\right|$$

such that $E\left|y^*_{it} - y_{it}\right|^2 \to 0$ and $y_{it} = y^*_{it}$ almost surely such that $\sup_i \sup_{t\geq 0} E\left[\left|y_{it}\right|^r\right] < \infty$. ■

Proof of Proposition 2. From Proposition 1 we know that y_{it} is stationary and that $\sup_i E\left[\left|y_{it}\right|^r\right] < \infty$. Define the σ-fields $\mathcal{A}^i_t \equiv \sigma\left(x_{it}, x_{it-1}, x_{it-2}, ...\right)$ and $\mathcal{B}^i_t \equiv \sigma\left(x_{it}, x_{it+1}, x_{it+2}, ...\right)$. Let $d = \dim(x_{it})$. Then for any Borel set $D \subset \mathbb{R}^d$ it follows that

$$P\left(x_{it+1} \in D | \mathcal{A}^i_t\right) = P\left(x_{it+1} \in D | x_{it}\right)$$

such that x_{it} is a Markov process. Since y_{it} is strictly stationary, x_{it} is also strictly stationary. By Theorem 7.11 of Kallenberg (1997) there exists an invariant measure

$$\vartheta\left(D\right) = P(x_{it} \in D). \tag{3}$$

Let

$$v_{it} = \left[\begin{array}{cc} 1 & \zeta'_0 \\ 0 & I_{d-1} \end{array}\right] \left[\begin{array}{c} \varepsilon_{it} \\ \eta_{it} \end{array}\right]$$

where $v_{it} = \left(v_{it,1}, v_{it,2}'\right)'$ is partitioned conformingly with $(\varepsilon_{it}, \eta_{it}')'$ and has density $p_{v_i}(v)$ with

$$p_{v_i}(v) = p_\varepsilon(v_1 - \zeta_0' v_2) p_\eta(v_2) .$$

Then, for $x = (x_y, x_z')' \in \mathbb{R}^d$ and partitioned conformingly with y_{it} and z_{it},

$$\begin{aligned}
\int |p_v(v - x) - p_v(v)| \, dv &= \int_v |p_\varepsilon(v_1 - x_y - \zeta_0'(v_2 - x_z)) p_\eta(v_2 - x_z) - p_\varepsilon(v_1 - \zeta_0' v_2) p_\eta(v_2)| \, dv \\
&\leq \int_v |p_\varepsilon(v_1 - x_y - \zeta_0'(v_2 - x_z)) - p_\varepsilon(v_1 - \zeta_0' v_2)| \, p_\eta(v_2 - x_z) \, dv \\
&\quad + \int_v |p_\eta(v_2 - x_z) - p_\eta(v_2)| \, p_\varepsilon(v_1 - \zeta_0' v_2) dv \\
&\leq C_\varepsilon (|x_y| + |\zeta_0' x_z|) \int p_\eta(v_2) dv_2 + C_\eta \|x_z\| .
\end{aligned} \tag{4}$$

Define $\theta = (\tau_0, \zeta_0' \Pi)'$ such that

$$x_{it} = \begin{bmatrix} \max\{0, \theta' x_{it-1} + v_{it,1} + \gamma_{i0}\} \\ \Pi z_{it-1} + v_{it,2} \end{bmatrix} .$$

For any Borel set $D \subset \mathbb{R}^d$ define

$$D^0 = \{(y,z) \in D | y = 0\} , \ D^y = \{(y,z) \in D | y > 0\}$$

where $D^0 \cap D^y = \emptyset$ and $D = D^0 \cup D^y$. Define $D_z^0 = \left\{z \in \mathbb{R}^{d-1} | \, (0, z) \in D^0\right\}$

$$\begin{aligned}
\Phi_v\left(D^0, x\right) &= \int_{v_2 + \Pi x_z \in D_z^0} \int_{v_1 \leq -\theta' x - \gamma_{i0}} p_v(v) dv_1 dv_2 \\
&= \int_{v_2 \in D_z^0} \int_{v_1 \leq 0} p_\varepsilon\left(v_1 - \zeta_0' v_2 - \tau_0' x_y - \gamma_{i0}\right) p_\eta\left(v_2 - \Pi x_z\right) dv_1 dv_2
\end{aligned}$$

and

$$\Phi_v^D(x) = \int_{(y,z) \in D^y} p_\varepsilon\left(y - \zeta_0' z - \tau_0' x_y - \gamma_{i0}\right) p_\eta\left(z - \Pi x_z\right) dy dz.$$

For any Borel set $D \subset \mathbb{R}^d$ and conditional on x_{it-1}, it follows that

$$P\left(x_{it} \in D | x_{it-1}\right) = P\left(x_{it} \in D^0 | x_{it-1}\right) + P\left(x_{it} \in D^y | x_{it-1}\right) = \Phi_v\left(D^0, x_{it-1}\right) + \Phi_v^D\left(x_{it-1}\right) . \tag{5}$$

From Athreya and Pantula (1986, p. 883) and exploiting the fact that x_{it} is a Markov process it follows that for any set $A \in \mathcal{A}_t^i$ and $B \in \mathcal{B}_{t+m}^i$, there exist Borel functions $g\left(x_{it+m}\right)$ and $h\left(x_{it}\right) : \mathbb{R}^d \to [0,1]$ such that ,

$$P(A \cap B) = E\left(g\left(x_{it+m}\right) h\left(x_{it}\right)\right)$$

and

$$P(A)P(B) = E\left(g\left(x_{it+m}\right)\right) E\left(h\left(x_{it}\right)\right).$$

Define $\mathcal{G} = \sigma\left(x_{it+m}\right), \mathcal{H} = \sigma\left(x_{it}\right).$ By Hall and Heyde (1980, Theorem A.5) and the fact that $g\left(x_{it+m}\right)$ is $\mathcal{G}$ measurable and $h\left(x_{it}\right)$ is $\mathcal{H}$ measurable it follows that

$$(6) \quad |P(A \cap B) - P(A)P(B)| = \left|E\left(g\left(x_{it+m}\right) h\left(x_{it}\right)\right) - E\left(g\left(x_{it+m}\right)\right) E\left(h\left(x_{it}\right)\right)\right| \leq 4\alpha\left(\mathcal{G}, \mathcal{H}\right)$$

where $\alpha\left(\mathcal{G}, \mathcal{H}\right) = \sup_{G \in \mathcal{G}, H \in \mathcal{H}} \left|P\left(G \cap H\right) - P(G)P(H)\right|$ does not depend on the functions $g()$ and $h()$. The remainder of the proof establishes a uniform bound for

$$(7) \quad |P(x_{it+m} \in D, x_{it} \in E) - P(x_{it+m} \in D)P(x_{it} \in E)|$$

for all $D \in \mathcal{R}^d$ and $E \in \mathcal{R}^d$ where $\mathcal{R}^d$ is the Borel field on $\mathbb{R}^d$. By Billingsley (1980, Theorem 20.1) the class of sets $\{x_{it} \in E\}$ for $E \in \mathcal{R}^d$ coincides with $\sigma\left(x_{it}\right)$ such that a uniform bound for (7) is also a bound for $\alpha\left(\mathcal{G}, \mathcal{H}\right)$ and thus for $\left|P(A \cap B) - P(A)P(B)\right|.$

The argument now follows a construction used in Gorodetskii (1977). Note that for some measurable function f one can write $x_{it+m} = f(\varepsilon_t^m, \eta_t^m, x_{it})$ where $\varepsilon_t^m = (\varepsilon_{it+1}, ..., \varepsilon_{it+m})'$ and $\eta_t^m = \left(\eta_{it+1}, ..., \eta_{it+m}\right)'$. Note that $f(\varepsilon_t^m, \eta_t^m, x)$ is independent of x_{it} for any fixed value $x \in \mathbb{R}^d$. Use the notation $x_{x,it+m}^m = f(\varepsilon_t^m, \eta_t^m, x)$ where $x_{x,it+m}^m = \left(y_{x,it+m}^m, z_{x,it+m}^{m\prime}\right)'$ to denote the process x_{it+m} with initial condition $x_{it} = x$. Let $u_{it+m}^m = y_{x,it+m}^m - y_{0,it+m}^m$. Let $x = (x_y, x_z')'$ denote the partition of the initial values of x_{it}. Then it follows that

$$\left|u_{x,it+m}^m\right| \leq \left|\tau_0\right|^m \left|x_y\right| + \left\|\zeta_0\right\| \textstyle\sum_{j=0}^{m-1} \left|\tau_0\right|^j \left\|\Pi^{m-j}\right\| \left\|x_z\right\|$$

and

$$\begin{aligned} \left\|x_{x,it+m}^m - x_{0,it+m}^m\right\| &\leq \left|y_{x,it+m}^m - y_{0,it+m}^m\right| + \left\|\zeta_0\right\| \left\|z_{x,it+m}^m - z_{0,it+m}^m\right\| \\ &\leq \left|\tau_0\right|^m \left|x_y\right| + 2\left\|\zeta_0\right\| \textstyle\sum_{j=0}^{m-1} \left|\tau_0\right|^j \left\|\Pi^{m-j}\right\| \left\|x_z\right\|. \end{aligned}$$

Here $\|\Pi\| = \left(\operatorname{tr} \Pi\Pi'\right)^{1/2}$ and $\Pi = TJT^{-1}$ is the Jordan decomposition (see Magnus and Neudecker, 1988, p. 17) where J is an upper diagonal matrix with the diagonal elements containing the eigenvalues of Π and T a nonsingular matrix. Then $\|\Pi^m\| = \|T\| \left\|T^{-1}\right\| \|J^m\|$ where

$$\|J^m\| = \left(\textstyle\sum_{j=1}^{d-1} \pi_j^{2m}\right)^{1/2} \leq \sqrt{(d-1)} \left|\bar{\pi}\right|^m$$

and $\bar{\pi} = \max_j \left|\pi_j\right|.$ For some constant $C_1 < \infty$ where C_1 does not depend on m, it follows that $\|\Pi\| \leq C_1 \left|\bar{\pi}\right|^m.$ Let $1 > a_0 > \max\left(\left|\tau_0\right|, \bar{\pi}\right)$ such that

$$\textstyle\sum_{j=0}^{m-1} \left|\tau_0\right|^j \left\|\Pi^{m-j}\right\| \leq C_1 \sum_{j=0}^{m-1} \left|\tau_0\right|^j \bar{\pi}^{m-j} = C_1 a_0^m \sum_{j=0}^{m-1} \left|\tau_0/a_0\right|^j \left|\bar{\pi}/a_0\right|^{m-j} = O(a_0^m).$$

It follows that $\left\| x^m_{x,it+m} - x^m_{0,it+m} \right\| \leq C a_0^m \left(|x_y| + \|x_z\| \right)$ for some constant C. Define $B = \left\{ x \in \mathbb{R}^d | \, \|x\| < \delta a_0^{-m} \right\}$ with complement B^c. Then,

$$P\left(x_{it+m} \in D, x_{it} \in E\right) = P\left(x_{it+m} \in D, x_{it} \in E \cap B\right) + P\left(x_{it+m} \in D, x_{it} \in E \cap B^c\right)$$

as well as

$$P\left(x_{it} \in E\right) P\left(x_{it+m} \in D\right) = P\left(x_{it} \in E\right) P\left(x_{it+m} \in D, x_{it} \in B\right) + P\left(x_{it} \in E\right) P\left(x_{it+m} \in D, x_{it} \in B^c\right)$$

such that

$$\begin{aligned} & |P(x_{it+m} \in D, x_{it} \in E) - P(x_{it+m} \in D) P(x_{it} \in E)| \\ \leq \; & |P(x_{it+m} \in D, x_{it} \in E \cap B) - P\left(x_{it+m} \in D, x_{it} \in B\right) P(x_{it} \in E)| \\ & + 2P\left(x_{it} \in B^c\right) \end{aligned} \tag{8}$$

Using (3) and denoting the joint distribution of $(\varepsilon_t^m, \eta_t^m)$ by $P_{\varepsilon,\eta}(.)$ we first write

$$\begin{aligned} P\left(x_{it+m} \in D, x_{it} \in E \cap B\right) &= \int_{B \cap E} \left(\int_{f(\varepsilon^m, \eta^m, x) \in D} P_{\varepsilon,\eta}\left(d\left(\varepsilon^m, \eta^m\right)\right) \right) \vartheta(dx) \\ &= \int_{B \cap E} P\left(f(\varepsilon_t^m, \eta_t^m, x) \in D\right) \vartheta(dx) \end{aligned}$$

where $P\left(f(\varepsilon_t^m, \eta_t^m, x) \in D\right) = \int_{f(\varepsilon^m, \eta^m, x) \in D} P_{\varepsilon,\eta}\left(d\left(\varepsilon^m, \eta^m\right)\right)$. Next note that

$$\int_{B \cap E} P\left(f(\varepsilon_t^m, \eta_t^m, 0) \in D\right) \vartheta(dx) = P\left(f(\varepsilon_t^m, \eta_t^m, 0) \in D\right) \int_{B \cap E} \vartheta(dx)$$

and

$$P(x_{it+m} \in D, x_{it} \in B) = \int_B P\left(f(\varepsilon_t^m, \eta_t^m, x) \in D\right) \vartheta(dx).$$

It then follows that

$$\begin{aligned} & P(x_{it+m} \in D, x_{it} \in E \cap B) - P(x_{it+m} \in D, x_{it} \in B) P(x_{it} \in E) \\ = \; & P(x_{it+m} \in D, x_{it} \in E \cap B) - P\left(f(\varepsilon_t^m, \eta_t^m, 0) \in D\right) P\left(x_{it} \in E\right) \\ & + P\left(f(\varepsilon_t^m, \eta_t^m, 0) \in D\right) P\left(x_{it} \in E\right) - P(x_{it+m} \in D, x_{it} \in B) P\left(x_{it} \in E\right) \\ \leq \; & P(x_{it+m} \in D, x_{it} \in E \cap B) - P\left(f(\varepsilon_t^m, \eta_t^m, 0) \in D\right) P\left(x_{it} \in E \cap B\right) \\ & + \left(P\left(f(\varepsilon_t^m, \eta_t^m, 0) \in D, x_{it} \in B\right) - P\left(x_{it+m} \in D, x_{it} \in B\right)\right) P\left(x_{it} \in E\right) \\ & + P\left(f(\varepsilon_t^m, \eta_t^m, 0) \in D, x_{it} \in B^c\right) P\left(x_{it} \in E\right) \\ \leq \; & \int_{B \cap E} \left| P\left(f(\varepsilon_t^m, \eta_t^m, x) \in D\right) - P\left(f(\varepsilon_t^m, \eta_t^m, 0) \in D\right) \right| \vartheta(dx) \\ & + \int_B \left| P\left(f(\varepsilon_t^m, \eta_t^m, x) \in D\right) - P\left(f(\varepsilon_t^m, \eta_t^m, 0) \in D\right) \right| \vartheta(dx) + P\left(x_{it} \in B^c\right) \\ \leq \; & 2 \int_B \left| P\left(f(\varepsilon_t^m, \eta_t^m, x) \in D\right) - P\left(f(\varepsilon_t^m, \eta_t^m, 0) \in D\right) \right| \vartheta(dx) + P\left(x_{it} \in B^c\right) \end{aligned} \tag{9}$$

as well as

$$
\begin{aligned}
(10) \qquad & P\left(x_{it+m} \in D, x_{it} \in B\right) P\left(x_{it} \in E\right) - P\left(x_{it+m} \in D, x_{it} \in E \cap B\right) \\
\leq \quad & \left(P(x_{it+m} \in D, x_{it} \in B) - P\left(f(\varepsilon_t^m, \eta_t^m, 0) \in D\right) P\left(x_{it} \in B\right)\right) P\left(x_{it} \in E\right) \\
& + P\left(f(\varepsilon_t^m, \eta_t^m, 0) \in D\right) P\left(x_{it} \in E\right) - P\left(x_{it+m} \in D, x_{it} \in E \cap B\right) \\
\leq \quad & \int_B \left|P\left(f(\varepsilon_t^m, \eta_t^m, x) \in D\right) - P\left(f(\varepsilon_t^m, \eta_t^m, 0) \in D\right)\right| \vartheta\left(dx\right) \\
& + P\left(f(\varepsilon_t^m, \eta_t^m, 0) \in D\right)\left(P\left(x_{it} \in E \cap B\right) + P\left(x_{it} \in E \cap B^c\right)\right) \\
& - P\left(x_{it+m} \in D, x_{it} \in E \cap B\right) \\
\leq \quad & \int_B \left|P\left(f(\varepsilon_t^m, \eta_t^m, x) \in D\right) - P\left(f(\varepsilon_t^m, \eta_t^m, 0) \in D\right)\right| \vartheta\left(dx\right) \\
& + \int_{B \cap E} \left|P\left(f(\varepsilon_t^m, \eta_t^m, x) \in D\right) - P\left(f(\varepsilon_t^m, \eta_t^m, 0) \in D\right)\right| \vartheta\left(dx\right) + P\left(x_{it} \in B^c\right) \\
\leq \quad & 2 \int_B \left|P\left(f(\varepsilon_t^m, \eta_t^m, x) \in D\right) - P\left(f(\varepsilon_t^m, \eta_t^m, 0) \in D\right)\right| \vartheta\left(dx\right) + P\left(x_{it} \in B^c\right)
\end{aligned}
$$

Now combine (8), (9) and (10) to conclude that

$$
\begin{aligned}
(11) \qquad & \left|P\left(x_{it+m} \in D, x_{it} \in E\right) - P\left(x_{it} \in E\right) P\left(x_{it+m} \in D\right)\right| \\
\leq \quad & 2 \int_B \left|P\left(f(\varepsilon_t^m, \eta_t^m, x) \in D\right) - P\left(f(\varepsilon_t^m, \eta_t^m, 0) \in D\right)\right| \vartheta\left(dx\right) + 3P\left(\|x_{it}\| \geq \delta a_0^{-m}\right)
\end{aligned}
$$

where $P\left(\|x_{it}\| \geq \delta a_0^{-m}\right) \leq a_0^{rm} \delta^{-r} E\left[\|x_{it}\|^r\right]$. Then

$$
\begin{aligned}
& \int_B \left|P\left(f(\varepsilon_t^m, \eta_t^m, x) \in D\right) - P\left(f(\varepsilon_t^m, \eta_t^m, 0) \in D\right)\right| \vartheta\left(dx\right) \\
\leq \quad & \sup_{x \in B} \left|P\left(x_{x,it+m}^m \in D\right) - P\left(x_{0,it+m}^m \in D\right)\right| \\
\leq \quad & \sup_{x \in B} E\left|P\left(x_{x,it+m}^m \in D | \varepsilon_t^{m-1}, \eta_t^{m-1}\right) - P\left(x_{0,it+m-1}^m \in D | \varepsilon_t^{m-1}, \eta_t^{m-1}\right)\right|
\end{aligned}
$$

where the last expectation operator runs over $\varepsilon_t^{m-1}, \eta_t^{m-1}$. Note that $z_{x,it+m-1} = \sum_{j=0}^{m-1} \Pi^j \eta_{it+m-j-1} + \Pi^m x_z$. From (5) it follows that

$$
\begin{aligned}
& \left|P\left(x_{x,it+m}^m \in D | \varepsilon_t^{m-1}, \eta_t^{m-1}\right) - P\left(x_{x_0,it+m-1}^m \in D | \varepsilon_t^{m-1}, \eta_t^{m-1}\right)\right| \\
\leq \quad & \left|\Phi_v(D^0, x_{x,it+m-1}^{m-1}) + \Phi_v(D^0, x_{0,it+m-1}^{m-1})\right| \\
& + \left|\Phi_v^D(x_{x,it+m-1}^{m-1}) - \Phi_v^D(x_{0,it+m-1}^{m-1})\right|
\end{aligned}
$$

where

$$
\begin{aligned}
(12)\quad & \left|\Phi_v^D(x^{m-1}_{x,it+m-1}) - \Phi_v^D(x^{m-1}_{0,it+m-1})\right| \\
\leq\ & \int_{(y,z)\in\mathbb{R}^d} \left|p_\varepsilon\left(y-\zeta_0'z-\tau_0 y^{m-1}_{x,it+m-1}-\gamma_{i0}\right)p_\eta\left(z-\Pi x^{m-1}_{x,it+m-1}\right)\right. \\
& \left. -p_\varepsilon\left(y-\zeta_0'z-\tau_0 y^{m-1}_{0,it+m-1}-\gamma_{i0}\right)p_\eta\left(z-\Pi x^{m-1}_{0,it+m-1}\right)\right| dydz \\
\leq\ & \int_{(y,z)\in\mathbb{R}^d} \left|p_\varepsilon\left(y-\zeta_0'z-\tau_0\left(y^{m-1}_{x,it+m-1}-y^{m-1}_{0,it+m-1}\right)-\gamma_{i0}\right)p_\eta\left(z-\Pi\left(z^{m-1}_{x,it+m-1}-z^{m-1}_{0,it+m-1}\right)\right)\right. \\
& \left. -p_v(y,z)\right| dydz \\
\leq\ & C_\varepsilon\left(\tau_0\left|y^{m-1}_{x,it+m-1}-y^{m-1}_{0,it+m-1}\right|+\|\zeta_0\|\left\|z^{m-1}_{x,it+m-1}-z^{m-1}_{0,it+m-1}\right\|\right)+C_\eta\left\|z^{m-1}_{x,it+m-1}-z^{m-1}_{0,it+m-1}\right\|.
\end{aligned}
$$

where the last inequality follows from (4). Then,

$$\left|y^{m-1}_{x,it+m-1}-y^{m-1}_{x_0,it+m-1}\right| \leq \tau_0^{m-1}|x_y| + \|\zeta_0\|\textstyle\sum_{j=0}^{m-2}|\tau_0|^j\left\|\Pi^{m-j}\right\|\|x_z\|$$

and

$$\left\|z^{m-1}_{x,it+m-1}-z^{m-1}_{0,it+m-1}\right\| \leq \left\|\Pi^{m-1}\right\|\|x_z\|$$

such that uniformly over all Borel sets $D\subset\mathbb{R}^d$

$$\left|\Phi_v^D(x^{m-1}_{x,it+m-1})-\Phi_v^D(x^{m-1}_{0,it+m-1})\right| \leq Ca_0^m\left(|x_y|+\|x_z\|\right)$$

for some constant C. The same inequality holds for $\left|\Phi_v(D^0,x^{m-1}_{x,it+m-1})+\Phi_v(D^0,x^{m-1}_{x_0,it+m-1})\right|$ which follows directly from the definition of $\Phi_v\left(D^0,.\right)$ and the first inequality in (12). Because the bounds do not depend on the errors ε_t^{m-1} and η_t^{m-1} it follows that

$$(13)\quad \sup_{x\in B} E\left|P\left(y^m_{x,it+m}\in D|\varepsilon_t^{m-1},\eta_t^{m-1}\right)-P\left(y^m_{x_0,it+m-1}\in D|\varepsilon_t^{m-1},\eta_t^{m-1}\right)\right| \leq 4C\delta$$

where we use the fact that $\sup_{x\in B}\left(|x_y|+\|x_z\|\right)\leq 2a_0^{-m}\delta$. Now choose $\varepsilon_a>0$ such that $a_0+\varepsilon_a<1$. Choose $\delta=(a_0+\varepsilon_a)^{m/r}$. Then,

$$(14)\quad P\left(|x_{it}|>\delta a_0^{-m}\right)\leq a_0^{rm}\delta^{-r}E\|x_{it}\|^r = a_0^r\left(a_0/\left(a_0+\varepsilon_a\right)\right)^m E\|x_{it}\|^r \to 0$$

and $\delta=\left((a_0+\varepsilon_a)^{1/r}\right)^m\to 0$ as $m\to\infty$. The result then follows from substituting upper bounds (13) and (14) in (11). ■

Proof of Proposition 3. By Propositions 1 and 2 it follows that x_{it} is strictly stationary and mixing with exponentially decaying coefficients and $\sup_i E\|x_{it}\|^r<\infty$ for $r>7+10q+12+\delta$ with $q\geq p/2+2$ and for some $\delta>0$. Consider $\xi(x)=\log\left(\int_{-\infty}^x(2\pi)^{-1/2}\exp\left(-1/2u^2\right)du\right)$, with the first five derivatives

$$\frac{d\xi(x)}{dx}=\frac{1}{2\sqrt{\pi}}\sqrt{2}\frac{e^{-\frac{1}{2}x^2}}{\frac{1}{2}\operatorname{erf}\left(\frac{1}{2}x\sqrt{2}\right)+\frac{1}{2}},$$

$$\frac{d^2\xi(x)}{(dx)^2} = -\frac{1}{2\pi}\frac{e^{2(-\frac{1}{2}x^2)}}{\left(\frac{1}{2}\,\mathrm{erf}\left(\frac{1}{2}x\sqrt{2}\right)+\frac{1}{2}\right)^2} - \frac{1}{2\sqrt{\pi}}x\sqrt{2}\frac{e^{-\frac{1}{2}x^2}}{\frac{1}{2}\,\mathrm{erf}\left(\frac{1}{2}x\sqrt{2}\right)+\frac{1}{2}}$$

$$\begin{aligned}
\frac{d^3\xi(x)}{(dx)^3} &= \frac{1}{\pi}x\frac{e^{-x^2}}{\left(\frac{1}{2}\,\mathrm{erf}\left(\frac{1}{2}x\sqrt{2}\right)+\frac{1}{2}\right)^2} + \frac{1}{2\pi}x\frac{e^{2(-\frac{1}{2}x^2)}}{\left(\frac{1}{2}\,\mathrm{erf}\left(\frac{1}{2}x\sqrt{2}\right)+\frac{1}{2}\right)^2} - \frac{1}{2\sqrt{\pi}}\sqrt{2}\frac{e^{-\frac{1}{2}x^2}}{\frac{1}{2}\,\mathrm{erf}\left(\frac{1}{2}x\sqrt{2}\right)+\frac{1}{2}} \\
&\quad + \frac{1}{2\sqrt{\pi}}x^2\sqrt{2}\frac{e^{-\frac{1}{2}x^2}}{\frac{1}{2}\,\mathrm{erf}\left(\frac{1}{2}x\sqrt{2}\right)+\frac{1}{2}} + \frac{1}{2\pi^{\frac{3}{2}}}\sqrt{2}e^{-x^2}\frac{e^{-\frac{1}{2}x^2}}{\left(\frac{1}{2}\,\mathrm{erf}\left(\frac{1}{2}x\sqrt{2}\right)+\frac{1}{2}\right)^3}
\end{aligned}$$

$$\begin{aligned}
\frac{d^4\xi(x)}{(dx)^4} &= \frac{3}{2\pi}\frac{e^{-x^2}}{\left(\frac{1}{2}\,\mathrm{erf}\left(\frac{1}{2}x\sqrt{2}\right)+\frac{1}{2}\right)^2} - \frac{3}{\pi}x^2\frac{e^{-x^2}}{\left(\frac{1}{2}\,\mathrm{erf}\left(\frac{1}{2}x\sqrt{2}\right)+\frac{1}{2}\right)^2} + \frac{1}{2\pi}\frac{e^{2(-\frac{1}{2}x^2)}}{\left(\frac{1}{2}\,\mathrm{erf}\left(\frac{1}{2}x\sqrt{2}\right)+\frac{1}{2}\right)^2} \\
&\quad - \frac{1}{2\pi}x^2\frac{e^{2(-\frac{1}{2}x^2)}}{\left(\frac{1}{2}\,\mathrm{erf}\left(\frac{1}{2}x\sqrt{2}\right)+\frac{1}{2}\right)^2} - \frac{3}{2\pi^2}e^{-x^2}\frac{e^{2(-\frac{1}{2}x^2)}}{\left(\frac{1}{2}\,\mathrm{erf}\left(\frac{1}{2}x\sqrt{2}\right)+\frac{1}{2}\right)^4} \\
&\quad + \frac{3}{2\sqrt{\pi}}x\sqrt{2}\frac{e^{-\frac{1}{2}x^2}}{\frac{1}{2}\,\mathrm{erf}\left(\frac{1}{2}x\sqrt{2}\right)+\frac{1}{2}} - \frac{1}{2\sqrt{\pi}}x^3\sqrt{2}\frac{e^{-\frac{1}{2}x^2}}{\frac{1}{2}\,\mathrm{erf}\left(\frac{1}{2}x\sqrt{2}\right)+\frac{1}{2}} \\
&\quad - \frac{3}{\pi^{\frac{3}{2}}}x\sqrt{2}e^{-x^2}\frac{e^{-\frac{1}{2}x^2}}{\left(\frac{1}{2}\,\mathrm{erf}\left(\frac{1}{2}x\sqrt{2}\right)+\frac{1}{2}\right)^3}
\end{aligned}$$

$$\begin{aligned}
\frac{d^5\xi(x)}{(dx)^5} &= -\frac{11}{\pi}x\frac{e^{-x^2}}{\left(\frac{1}{2}\,\mathrm{erf}\left(\frac{1}{2}x\sqrt{2}\right)+\frac{1}{2}\right)^2} + \frac{7}{\pi}x^3\frac{e^{-x^2}}{\left(\frac{1}{2}\,\mathrm{erf}\left(\frac{1}{2}x\sqrt{2}\right)+\frac{1}{2}\right)^2} \\
&\quad + \frac{6}{\pi^2}x\frac{e^{2(-x^2)}}{\left(\frac{1}{2}\,\mathrm{erf}\left(\frac{1}{2}x\sqrt{2}\right)+\frac{1}{2}\right)^4} - \frac{3}{2\pi}x\frac{e^{2(-\frac{1}{2}x^2)}}{\left(\frac{1}{2}\,\mathrm{erf}\left(\frac{1}{2}x\sqrt{2}\right)+\frac{1}{2}\right)^2} \\
&\quad + \frac{1}{2\pi}x^3\frac{e^{2(-\frac{1}{2}x^2)}}{\left(\frac{1}{2}\,\mathrm{erf}\left(\frac{1}{2}x\sqrt{2}\right)+\frac{1}{2}\right)^2} + \frac{9}{\pi^2}xe^{-x^2}\frac{e^{2(-\frac{1}{2}x^2)}}{\left(\frac{1}{2}\,\mathrm{erf}\left(\frac{1}{2}x\sqrt{2}\right)+\frac{1}{2}\right)^4} \\
&\quad + \frac{3}{2\sqrt{\pi}}\sqrt{2}\frac{e^{-\frac{1}{2}x^2}}{\frac{1}{2}\,\mathrm{erf}\left(\frac{1}{2}x\sqrt{2}\right)+\frac{1}{2}} - \frac{3}{\sqrt{\pi}}x^2\sqrt{2}\frac{e^{-\frac{1}{2}x^2}}{\frac{1}{2}\,\mathrm{erf}\left(\frac{1}{2}x\sqrt{2}\right)+\frac{1}{2}} \\
&\quad + \frac{1}{2\sqrt{\pi}}x^4\sqrt{2}\frac{e^{-\frac{1}{2}x^2}}{\frac{1}{2}\,\mathrm{erf}\left(\frac{1}{2}x\sqrt{2}\right)+\frac{1}{2}} - \frac{5}{\pi^{\frac{3}{2}}}\sqrt{2}e^{-x^2}\frac{e^{-\frac{1}{2}x^2}}{\left(\frac{1}{2}\,\mathrm{erf}\left(\frac{1}{2}x\sqrt{2}\right)+\frac{1}{2}\right)^3} \\
&\quad + \frac{3}{\pi^{\frac{5}{2}}}\sqrt{2}e^{2(-x^2)}\frac{e^{-\frac{1}{2}x^2}}{\left(\frac{1}{2}\,\mathrm{erf}\left(\frac{1}{2}x\sqrt{2}\right)+\frac{1}{2}\right)^5} + \frac{25}{2\pi^{\frac{3}{2}}}x^2\sqrt{2}e^{-x^2}\frac{e^{-\frac{1}{2}x^2}}{\left(\frac{1}{2}\,\mathrm{erf}\left(\frac{1}{2}x\sqrt{2}\right)+\frac{1}{2}\right)^3}
\end{aligned}$$

where

$$\frac{1}{2}\,\mathrm{erf}\left(\frac{1}{2}x\sqrt{2}\right)+\frac{1}{2} = \Lambda(x).$$

Therefore, $\frac{d^{|v|}\xi(x)}{(dx)^{|v|}}$ is bounded except for $x \to -\infty$. We note that $\frac{e^{-x^2/2}}{\Lambda(x)} = O(-x)$ as $x \to -\infty$ which implies that

$$\frac{d^{|v|}\xi(x)}{(dx)^{|v|}} = O(|x|^{|v|}) \text{ as } x \to -\infty$$

and bounded elsewhere. We thus have the bound $\left| \frac{d^{|v|}\xi(x)}{(dx)^{|v|}} \right| \leq C\,|x|^{|v|}$ for some constant C. This now implies that

$$\sup_{\beta \in \Phi} \|D^v \psi(x_{it}, \beta)\| \leq C \sup_{\beta \in \Phi} \left| x_{it}' \tilde{\beta} \right|^{|v|} \leq C\,\|x_{it}\|^{|v|} \sup_{\beta \in \Phi} \left| \tilde{\beta} \right|^{|v|}$$

where $\tilde{\beta} = (1, \phi')'$. The assertion then follows from $\sup_i E\,\|x_{it}\|^{|v|+10q+12+\delta} < \infty$ since $\sup_{\phi \in \Phi} \left| \tilde{\beta} \right|^{|v|}$ is bounded when Φ is compact. By the chain rule of differentiation and the mean value theorem we have

$$|D^v \psi(x_{it}, \phi_1) - D^v \psi(x_{it}, \phi_2)| \leq C \left| x_{it}' \tilde{\beta}^* \right|^{|v|+1} \|x_{it}\|\,\|\phi_1 - \phi_2\|$$

where $\left\| \tilde{\beta}^* - \tilde{\beta}_1 \right\| \leq \left\| \tilde{\beta}_1 - \tilde{\beta}_2 \right\|$ and $\left\| \tilde{\beta}^* - \tilde{\beta}_2 \right\| \leq \left\| \tilde{\beta}_1 - \tilde{\beta}_2 \right\|$ such that the result follows from

$$\sup_i E\,\|x_{it}\|^{|v|+2+10q+12+\delta} < \infty$$

which holds by Propositions 1. ■

References

[1] Athreya, K.B and S.G. Pantula (1986): "Mixing Properties of Harris Chains and Autoregressive Processes", *J. Appl. Prob.* 23, 880-892.

[2] Billingsley, P. (1980): "*Probability and Measure*", 3rd edition, Wiley, New York.

[3] De Jong, R. and A. Herrera (2005): "Dynamic censored regression and the Open Market Desk reaction function", manuscript.

[4] De Jong, R. and T. Woutersen (2005): "Dynamic time series binary choice", manuscript.

[5] Gorodetskii, V.V. (1977): "On the Strong Mixing Property for Linear Sequences", *Theory of Probability and its Applications,* 22, 411-413.

[6] Hahn, J. and G. Kuersteiner (2004): "Bias Reduction for Dynamic Nonlinear Panel Models", *manuscript.*

[7] Hall, P., and C. Heyde, (1980): "*Martingale Limit Theory and its Applications*", Academic Press, New York.

[8] Kallenberg, P. (1997): *Probability and its Applications.* Springer Verlag, New York.

[9] Neyman, J., and E. Scott (1948): "Consistent Estimates Based on Partially Consistent Observations", *Econometrica* 16, pp. 1 -31.

[10] Potscher, B.M. and I.R. Prucha (1991): *Dynamic Nonlinear Econometric Models. Asymptotic Theory.* Springer Verlag, Berlin, Heidelberg, New York.

[11] Tobin, J. (1958): "Estimation of Relationships for Limited Dependent Variables.", *Econometrica* 26, 24-36.