Appendices to DYNAMIC QUANTILE MODELS

C. GOURIEROUX * and J. JASIAK **

First version: November, 2005; Revised May, 2007

[1]CREST, CEPREMAP and University of Toronto.
[2]York University.
The authors gratefully acknowledge financial support of Natural Sciences and Engineering Council (NSERC), Canada, and of the Chair AXA: "Large Risks in Insurance".

Appendix 1
Proposition 1

The asymptotic properties of the information-based estimator are based on standard results on empirical processes of stationary mixing sequences [see e.g. Arcones, Yu (1994), VanderVaart, Wellner (1996)] and on the almost sure uniform convergence of the kernel estimator of density function for serially dependent observations [see e.g. Tenreiro (1995), Th. 2.2.6, or Hansen (2006)] The latter condition [see Assumptions A.5, A.16 on the tail behavior and condition A.18 on the bandwidth] is required to prove the uniform convergence of the standardized criterion function and to derive the (a.s) convergence of the estimator.

The set of sufficient regularity conditions is given below. Next, the asymptotic expansions are provided, to justify the asymptotic normality and efficiency of the estimators. Note that we are interested in the asymptotic properties of the parameter estimator and not in the asymptotic behavior of the kernel estimator itself. The asymptotic efficiency of the parameter estimator can be reached even without an optimal bandwidth selection. This explains why Assumption A.17 can eliminate the asymptotic bias and the optimal nonparametric rate.

A.1 Regularity Conditions

Below, we provide a set of sufficient conditions to derive the asymptotic behavior of the information-based estimator. These assumptions can be weakened, for example, to accommodate non-Markov processes, more conditioning variables or to account for the fat tails of the distribution (see the discussion of the trimmed estimator in section 3.3). As usual, $||A||$ denotes the Frobenius norm of a matrix, which reduces to the Euclidean norm when A is a vector. C^m denotes the space of functions f, which are continuously differentiable up to order m, and $||D^m f||_\infty = ||d^m f(y)/dy^m||_\infty$.

i) Assumptions on the process

Assumption A.1 : The process (y_t) is a univariate continuous Markov process, strictly stationary and geometric strong mixing.

Assumption A.2: $x_t = y_{t-1}$.

Assumption A.3 : The stationary density f of Y_t is in class C^m, for some $m \in N, m \geq 2$, such that $||f||_\infty < \infty$ and $||D^m f||_\infty < \infty$.

Assumption A.4 : The stationary density f_t of Y_t and f_{t_1,t_2} of (Y_{t_1}, Y_{t_2}) are such that $||f_t||_\infty < \infty$, and $sup_{t_1<t_2}||f_{t_1,t_2}||_\infty < \infty$, respectively. Moreover, for $t_1 < t_2 < t_3 < t_4$, the stationary density f_{t_1,t_2,t_3,t_4} of $(Y_{t_1}, Y_{t_2}, Y_{t_3}, Y_{t_4})$ is such that:

$$sup_{t_1<t_2<t_3<t_4}||f_{t_1,t_2,t_3,t_4}||_\infty < \infty.$$

Assumption A.5 : There exists a positive constant $c \geq 2$, such that:
i) $E(|Y_t|^{2c}) < \infty$;
ii) $|||y|^2 f_t(y)||_\infty < \infty$.

ii) Assumptions on the model

Assumption A.6 : The dynamic quantile model is well-specified, with θ_0 as the true value of the parameter.

Assumption A.7 : The parameter space is compact and θ_0 is in the interior of the parameter space.

Assumption A.8 : θ_0 is identifiable.

Assumption A.9 : The dynamic quantile model is DAQ.

The functions $a_k(x, \alpha_k),\ k = 0, ..., K$ are twice continuously differentiable with respect to $x, \alpha_k,\ k = 0, ..., K$.

The baseline quantile functions $Q_k(., \beta_k),\ k = 1, ..., K$ are twice continuously differentiable with respect to $\beta_k,\ k = 1, ..., K$.

Assumption A.10 : Let us denote $f(y|x;\theta)$ the conditional density of y_t given x_t. The mapping $x \to sup_\theta \int f(y|x;\theta)\log(f(y|x;\theta)/f(y|x;\theta_0))dyf(x)$ is bounded

Assumption A.11: There exists $\delta, \gamma > 1$, $\tau > 0$ such that

$$\begin{aligned}
E\left(||\frac{\partial \log f(Y_t|X_t;\theta_0)}{\partial\theta}||^\tau\right) &< \infty, \\
E\left(sup_\theta||\frac{\partial \log f(Y_t|X_t;\theta_0)}{\partial\theta}||^\delta\right) &< \infty, \\
E\left(sup_\theta||\frac{\partial^2 \log f(Y_t|X_t;\theta_0)}{\partial\theta\partial\theta'}||^\gamma\right) &< \infty.
\end{aligned}$$

Assumption A.12: The mapping

$$x \rightarrow E\left(sup_\theta || \frac{\partial \log f(Y_t|X_t;\theta_0)}{\partial \theta} ||^2 | X_0 = x \right) f(x)$$

is bounded

Assumption A.13: The functions

$$\begin{aligned}
\theta &\rightarrow E\left(\frac{\partial \log f(Y_t|X_t;\theta_0)}{\partial \theta} \right), \\
\theta &\rightarrow E\left(\frac{\partial^2 \log f(Y_t|X_t;\theta_0)}{\partial \theta \partial \theta'} \right), \\
\theta &\rightarrow E\left(\frac{\partial \log f(Y_t|X_t;\theta_0)}{\partial \theta} | X_t = x \right),
\end{aligned}$$

are continuous on the interior of the parameter space, for any x.

iii) Assumptions on the kernels and bandwidths

Assumption A.14 : $K^* = K, h_T^* = h_T$, that is, we consider a product kernel.

Assumption A.15 : The kernel K is a Parzen kernel of order m, that is :
i) $\int K(y)dy = 1$;
ii) K is bounded, $\lim_{|y| \rightarrow \infty} |y| K(y) = 0$, $\int |K(y)| dy < \infty$, $\int K(y)^2 dy < \infty$;
iii) $\int y^l K(y) dy = 0$, for any $l \in \mathcal{N}$, such that $l < m$, and $\int K(y)|y|^m dy < \infty$.

Assumption A.16 : The kernel K is differentiable, $||\frac{dK(u)}{du}||_\infty < \infty$ and there exist positive constants c_0, c_1, such that $|\frac{dK(u)}{du}| < c_0 |u|^c$ for $|u| > c_1$, where the constant c is defined in Assumption A.5.

Assumption A.17 : The bandwidth h_T is such that : $T h_T^{2+2m} \rightarrow 0$, as $T \rightarrow \infty$.

Assumption A.18 : The bandwidth is such that $T^\alpha h_T^2 / \log T \rightarrow \infty$, as $T \rightarrow \infty$ for some $\alpha < 0.5$.

A.2 Asymptotic Expansions

The estimator is a solution to the following optimization:

$$
\begin{aligned}
\hat{\theta}_T &= \arg\min_\theta \sum_{t=1}^{T} \left\{ \int f(y|x_t;\theta) \log f(y|x_t;\theta) dy \right. \\
&\qquad \left. - \int f(y|x_t;\theta) \log \hat{f}_{0T}(y|x_t) dy \right\} \\
&= \arg\min_\theta \sum_{t=1}^{T} \left\{ \int f_t(y;\theta) \log f_t(y;\theta) dy - \int f_t(y;\theta) \log \hat{f}_{0,T,t}(y) dy \right\},
\end{aligned}
$$

say.

i) First-order conditions

They are given by :

$$
\sum_{t=1}^{T} \left\{ \int \frac{\partial f_t}{\partial \theta}(y;\hat{\theta}_T) \log f_t(y;\hat{\theta}_T) dy + \int \frac{\partial f_t}{\partial \theta}(y;\hat{\theta}_T) dy - \int \frac{\partial f_t}{\partial \theta}(y;\hat{\theta}_T) \log \hat{f}_{0,T,t}(y) dy \right\}
$$

$= 0,$

or :

$$
\sum_{t=1}^{T} \left\{ \int \frac{\partial f_t}{\partial \theta}(y;\hat{\theta}_T) \log f_t(y;\hat{\theta}_T) dy - \int \frac{\partial f_t}{\partial \theta}(y;\hat{\theta}_T) \log \hat{f}_{0,T,t}(y) dy \right\} = 0,
$$

since : $\displaystyle\int \frac{\partial f_t}{\partial \theta}(y;\hat{\theta}_T) dy = 0.$

ii) Expansion of the first-order conditions

Let us consider the expansion when T is large and $\hat{\theta}_T$ is in a neighbourhood of θ_0. We get :

$$
\sum_{t=1}^{T} \left\{ \int \left[\frac{\partial f_t}{\partial \theta}(y;\theta_0) + \frac{\partial^2 f_t}{\partial \theta \partial \theta'}(y;\theta_0)(\hat{\theta}_T - \theta_0) \right] \log[f_t(y;\theta_0) + \frac{\partial f_t}{\partial \theta'}(y;\theta_0)(\hat{\theta}_T - \theta_0)] dy \right\}
$$

$$- \int \left[\frac{\partial f_t}{\partial \theta}(y;\theta_0) + \frac{\partial^2 f_t}{\partial \theta \partial \theta'}(y;\theta_0)(\hat{\theta}_T - \theta_0) \right] \log \left[f_t(y;\theta_0) + \hat{f}_{0,T,t}(y) - f_t(y;\theta_0) \right] dy \Big\} \simeq 0,$$

or :

$$\sum_{t=1}^{T} \int \frac{1}{f_t(y;\theta_0)} \frac{\partial f_t}{\partial \theta}(y;\theta_0) \frac{\partial f_t}{\partial \theta'}(y;\theta_0) dy (\hat{\theta}_T - \theta_0)$$

$$- \sum_{t=1}^{T} \int \frac{\partial f_t}{\partial \theta}(y;\theta_0) \frac{1}{f_t(y;\theta_0)} [\hat{f}_{0,T,t}(y) - f_t(y;\theta_0)] dy \simeq 0.$$

We deduce that :

$$\begin{aligned}
\sqrt{T}(\hat{\theta}_T - \theta_0) \quad & \simeq \left[\frac{1}{T} \sum_{t=1}^{T} E_t \left[\frac{\partial \log f_t(y;\theta_0)}{\partial \theta} \frac{\partial \log f_t(y;\theta_0)}{\partial \theta'} \right] \right]^{-1} \\
& \sqrt{T} \sum_{t=1}^{T} \int \frac{\partial \log f_t(y;\theta_0)}{\partial \theta} \left[\hat{f}_{0,T,t}(y) - f_t(y;\theta_0) \right] dy \\
& \simeq \left(E \left[\frac{\partial \log f_t(y;\theta_0)}{\partial \theta} \frac{\partial \log f_t(y;\theta_0)}{\partial \theta'} \right] \right)^{-1} \\
& \frac{1}{\sqrt{T}} \sum_{t=1}^{T} \int \frac{\partial \log f_t(y;\theta_0)}{\partial \theta} [\hat{f}_{0,T,t}(y) - f_t(y;\theta_0)] dy.
\end{aligned}$$

The result follows from the asymptotic properties of the kernel estimator of the conditional density. Indeed, we get :

$$1/\sqrt{T} \sum_{t=1}^{T} \int \frac{\partial \log f_t(y;\theta_0)}{\partial \theta} \left[\hat{f}_{0,T,t}(y) - f_t(y;\theta_0) \right] dy$$

$$\rightsquigarrow N \left(0, E \left[\frac{\partial \log f_t(y;\theta_0)}{\partial \theta} \frac{\partial \log f_t(y;\theta_0)}{\partial \theta'} \right] \right).$$

The asymptotic bias of the kernel estimator of density has been eliminated by the appropriate choice of the bandwidth (see Assumption A.17). The parametric rate of convergence is due to the integral expression of the asymptotic equivalent.

Appendix 2
Asymptotic Expansion of a sample moment condition

Let us assume that the instrumental variable is a function of variable $X : Z_t = Z(X_t)$, say, and consider the expansion of the sample moment condition. We get:

$$
\begin{aligned}
&\frac{1}{\sqrt{T}}\sum_{t=1}^{T}\Big\{Z(X_t)Y_t g[\hat{F}_T(Y_t|X_t)] - E[Z(X_t)Y_t g[F_0(Y_t|X_t)]]\Big\} \\
\approx\;& \frac{1}{\sqrt{T}}\sum_{t=1}^{T}\{Z(X_t)Y_t g[F_0(Y_t|X_t)] - E[Z(X_t)Y_t g[F_0(Y_t|X_t)]]\} \\
&+\frac{1}{\sqrt{T}}\sum_{t=1}^{T} Z(X_t)Y_t\frac{dg}{du}[F_0(Y_t|X_t)][\hat{F}_T(Y_t|X_t) - F_0(Y_t|X_t)] \\
\approx\;& \int\int Z(x)yg[F_0(y|x)]d\sqrt{T}[\hat{G}_T(y,x) - G_0(y,x)] \\
&+\int\int Z(x)y\frac{dg}{du}(F_0(y|x))\sqrt{T}[\hat{F}_T(y|x) - F_0(y|x)]dG_0(y,x),
\end{aligned}
$$

where $G_0(y,x)$ is the joint cdf of (Y_t, X_t) and $\hat{G}_T$ its sample counterpart.

Similarly, let us denote by $\hat{G}_T(x)$ [resp $G_0(x)$] the sample cdf of X_t [resp. the unconditional cdf of X_t]. We get:

$$
\begin{aligned}
&\frac{1}{\sqrt{T}}\sum_{t=1}^{T}\Big\{Z(X_t)Y_t g[\hat{F}_T(Y_t|X_t)] - E[Z(X_t)Y_t g[F_0(Y_t|X_t)]]\Big\}\\
\approx\ & \int\int Z(x)yg[F_0(y|x)]d[\sqrt{T}[\hat{F}_T(y|x)-F_0(y|x)]G_0(x)+F_0(y|x)\sqrt{T}[\hat{G}_T(x)-G_0(x)]]\\
&+\int\int Z(x)y\frac{dg}{du}[F_0(y|x)]\sqrt{T}[\hat{F}_T(y|x)-F_0(y|x)]d[G_0(x)F_0(y|x)]\\
\approx\ & \int Z(x)[\int yg[F_0(y|x)]d[\sqrt{T}[\hat{F}_T(y|x)-F_0(y|x)]]dG_0(x)\\
&+\int Z(x)[\int yg[F_0(y|x)]dF_0(y|x)]d\sqrt{T}[\hat{G}_T(x)-G_0(x)]\\
&+\int Z(x)[\int y\frac{dg}{du}[F_0(y|x)]\sqrt{T}[\hat{F}_T(y|x)-F_0(y|x)]dF_0(y|x)]dG_0(x).
\end{aligned}
$$

Thus, the asymptotic behavior of the sample moment condition depends on the joint asymptotic behavior of the processes ($\sqrt{T}[\hat{G}_T(x)-G_0(x)]$, $\sqrt{T}[\hat{F}_T(y|x)-F_0(y|x)]$, which are indexed by x and y. Loosely speaking, if the observations $(X_t, Y_t), t = 1, .., T$ were iid, we could apply standard functional limit theorems and the asymptotic independence between processes ($\sqrt{T}[\hat{G}_T(x) - G_0(x)]$ and $\sqrt{T}[\hat{F}_T(y|x) - F_0(y|x)]$, for any x, to deduce from the expansion above the asymptotic normality of the sample moment condition and a rather simple expression of the asymptotic variance-covariance matrix. When $X_t = Y_{t-1}$, the asymptotic normality still holds under reasonable regularity conditions, but the expression of the asymptotic variance has to take into account serial dependence and is rather cumbersome.