2001s-15

# Properties of Estimates of Daily GARCH Parameters Based on Intra-Day Observations

John W. Galbraith, Victoria Zinde-Walsh

Série Scientifique Scientific Series



Montréal Février 2001

#### CIRANO

Le CIRANO est un organisme sans but lucratif constitué en vertu de la Loi des compagnies du Québec. Le financement de son infrastructure et de ses activités de recherche provient des cotisations de ses organisationsmembres, d'une subvention d'infrastructure du ministère de la Recherche, de la Science et de la Technologie, de même que des subventions et mandats obtenus par ses équipes de recherche.

CIRANO is a private non-profit organization incorporated under the Québec Companies Act. Its infrastructure and research activities are funded through fees paid by member organizations, an infrastructure grant from the Ministère de la Recherche, de la Science et de la Technologie, and grants and research mandates obtained by its research teams.

#### Les organisations-partenaires / The Partner Organizations

•École des Hautes Études Commerciales •École Polytechnique •Université Concordia •Université de Montréal •Université du Ouébec à Montréal •Université Laval •Université McGill •MEO •MRST •Alcan Aluminium Ltée •AXA Canada •Banque du Canada •Banque Laurentienne du Canada •Banque Nationale du Canada •Banque Royale du Canada •Bell Ouébec •Bombardier •Bourse de Montréal •Développement des ressources humaines Canada (DRHC) •Fédération des caisses populaires Desjardins de Montréal et de l'Ouest-du-Québec •Hydro-Québec •Imasco •Industrie Canada •Pratt & Whitney Canada Inc. •Raymond Chabot Grant Thornton •Ville de Montréal

© 2001 John W. Galbraith et Victoria Zinde-Walsh. Tous droits réservés. All rights reserved. Reproduction partielle permise avec citation du document source, incluant la notice ©. Short sections may be quoted without explicit permission, if full credit, including © notice, is given to the source.

Ce document est publié dans l'intention de rendre accessibles les résultats préliminaires de la recherche effectuée au CIRANO, afin de susciter des échanges et des suggestions. Les idées et les opinions émises sont sous l'unique responsabilité des auteurs, et ne représentent pas nécessairement les positions du CIRANO ou de ses partenaires. *This paper presents preliminary research carried out at CIRANO and aims at encouraging discussion and comment. The observations and viewpoints expressed are the sole responsibility of the authors. They do not necessarily represent positions of CIRANO or its partners.* 

#### ISSN 1198-8177

# **Properties of Estimates of Daily GARCH Parameters Based on Intra-Day Observations**<sup>\*</sup>

John W. Galbraith<sup> $\dagger$ </sup>, Victoria Zinde-Walsh<sup> $\ddagger$ </sup>

### Résumé / Abstract

Nous considérons les estimés des paramètres des modèles GARCH pour les rendements financiers journaliers, qui sont obtenus à l'aide des données intrajour (haute fréquence) pour estimer la volatilité journalière. Deux bases potentielles sont evaluées. La première est fondée sur l'aggrégation des estimés quasi-vraisemblance-maximale, en profitant des résultats de Drost et Nijman (1993). L'autre utilise la volatilité integrée de Andersen et Bollerslev (1998), et obtient les coefficients d'un modèle estimé par LAD ou MCO; la première méthode résiste mieux à la possibilité de non-existence des moments de l'erreur en estimation de volatilité. En particulier, nous considérons l'estimation par approximation ARCH, et nous obtenons les paramètres par une méthode liée à celle de Galbraith et Zinde-Walsh (1997) pour les processus ARMA. Nous offrons des résultats provenant des simulations sur la performance des méthodes en échantillons finis, et nous décrivons les atouts relatifs à l'estimation standard de quasi-VM basée uniquement sur les données journalières.

We consider estimates of the parameters of GARCH models of daily financial returns, obtained using intra-day (high-frequency) returns data to estimate the daily conditional volatility.Two potential bases for estimation are considered. One uses aggregation of high-frequency Quasi- ML estimates, using aggregation results of Drost and Nijman (1993). The other uses the integrated volatility of Andersen and Bollerslev (1998), and obtains coefficients from a model estimated by LAD or OLS, in the former case providing consistency and asymptotic normality in some cases where moments of the volatility estimation error may not exist. In particular, we consider estimation in this way of an ARCH approximation, and obtain GARCH parameters by a method related to that of Galbraith and Zinde-Walsh (1997) for ARMA processes. We offer some simulation evidence on small-sample performance, and characterize the gains relative to standard quasi-ML estimates based on daily data alone.

<sup>&</sup>lt;sup>\*</sup> Corresponding Author: John W. Galbraith, Department of Economics, McGill University, 888 Sherbrooke Street West, Montréal, Qc, Canada H3A 2T7 Tel.: (514) 985-4008 Fax: (514) 985-4039 email: galbraij@cirano.qc.ca The authors thank the Fonds pour la formation de chercheurs et l'aide à la recherche (Québec) and the Social Sciences and Humanities Research Council of Canada for financial support of this research, and CIRANO for research facilities. We are grateful to Nour Meddahi, Tom McCurdy and Éric Renault, and to participants in the Econometric Society World Congress 2000 and Canadian Econometric Study Group 2000 meeting, for valuable comments and discussions.

<sup>&</sup>lt;sup>†</sup> McGill University and CIRANO

<sup>&</sup>lt;sup>‡</sup> McGill University

- Mots Clés: GARCH, données haute fréquence, volatilité intégrée, LAD
- Keywords: GARCH, high frequency data, integrated volatility, LAD

**JEL:** C22

#### 1. Introduction

GARCH models are widely used for forecasting and characterizing the conditional volatility of economic and (particularly) financial time series. Since the original contributions of Engle (1982) and Bollerslev (1986), the models have been estimated by Maximum Likelihood (or quasi-Maximum Likelihood, 'QML') methods on observations at the frequency of interest. In the case of asset returns, the frequency of interest is often the daily fluctuation.

Financial data are often recorded at frequencies much higher than the daily. Even where our interest lies in volatility at the daily frequency, these data contain information which may be used to improve our estimates of models at the daily frequency. Of course, following Andersen and Bollerslev (1998), higher-frequency data may also be used to estimate the daily volatility directly.

The present paper considers two possible strategies for estimation of daily GARCH models which use information about higher-frequency fluctuations. The first uses the known aggregation relations (Drost and Nijman, 1993) linking the parameters of GARCH models of high-frequency and corresponding low-frequency observations. When such estimates are based on QML estimates for the high-frequency data, however, relatively stringent conditions are required, which may not be met in (for example) asset-return data.

The second potential strategy is to use the observation of Andersen and Bollerslev (1998) that the volatility of low-frequency asset returns may be estimated by the sum of squared high-frequency returns. While the resulting estimate may be used directly to characterize the process as in Andersen and Bollerslev or Andersen et al. (1999), it is also possible to use the sequence of low- (daily-) frequency estimated volatilities to obtain estimates of conditional volatility models such as GARCH models, explicitly allowing for estimation error in the estimated daily volatility. The resulting models may be estimated by a variety of techniques (including LS); by using the Least Absolute Deviations (LAD) estimator, it is possible to obtain consistent and asymptotically normal estimates under quite general conditions (in particular, without requiring the existence of moments of the returns). We are then able to obtain estimates of GARCH parameters using an estimator related to that of Galbraith and Zinde-Walsh (1997) for ARMA models.

In section 2 we describe the models and estimators to be considered and give some relevant definitions and notation. Section 3 provides several asymptotic results, while section 4 presents simulation evidence on the finite- sample performance of regression estimators relative to that of standard GARCH estimates based on the daily observations alone.

#### 2. GARCH model estimation using higher-frequency data

#### 2.1 Processes and notation

We begin by establishing notation for the processes of interest. Consider a driftless

diffusion process  $\{X_t\}$  such that

$$X_t = X_0 + \int_0^t \sigma_s X_s dW_s,$$

where  $\{W_t\}$  is a Brownian motion process and  $\sigma_s^2$  is the instantaneous conditional variance. This is a special case of the structure used by, e.g., Nelson (1992), Nelson and Foster (1994).

The process is sampled discretely at an interval of time  $\ell$  (e.g., each minute). We are interested in volatility at a lower-frequency sampling, with sampling interval  $h\ell$  (e.g., daily), so that there are h high-frequency observations per low-frequency observation. Define one unit of time as a period of length  $\ell$ .

We index the full set of observations by  $\tau$  and the lower-frequency observations by t, so that  $t = \{h, 2h, \ldots, hT\}$ . The size of the sample of low-frequency observations is therefore T, and of the full set of high-frequency observations is hT. Following Andersen and Bollerslev (1998), estimate the conditional volatility at t as the estimated conditional variance

$$\hat{\sigma}_t^2 = \sum_{j=(i-1)h+1}^{ih} r_j^2, \qquad (2.1.0)$$

with  $r_j^2 = (x_j - x_{j-1})^2$ ,  $x_j$  indicating the discretely-sampled observations on the process X. See Andersen and Bollerslev on convergence of  $\hat{\sigma}_t^2$  to  $\int_{t-1}^t \sigma_s^2 ds$ .

Now consider ARCH and GARCH models at the lower-frequency observations:

$$\sigma_t^2 = \omega + \sum_{i=1}^q \alpha_i \varepsilon_{t-i}^2, \qquad (2.1.1)$$

$$\sigma_t^2 = \omega + \sum_{i=1}^q \alpha_i \varepsilon_{t-i}^2 + \sum_{i=1}^p \beta_i \sigma_{t-i}^2, \qquad (2.1.2)$$

where  $\varepsilon_t = y_t - \mu_t$  for a process  $y_t$  with conditional mean  $\mu_t$ , or in the driftless case  $\varepsilon_t = x_t$ . So  $E(\varepsilon_t^2 | \psi_{t-i}) \equiv \sigma_t^2$ . Models in the form (2.1.1), (2.1.2) are directly estimable if, as in Andersen and Bollerslev, we have measurements of  $\sigma_t^2$ . We return to this point in Section 2.3 below.

Finally, we will refer below to the definitions of Strong, Semi-strong and Weak GARCH given in Drost and Nijman (1993). In strong GARCH,  $\{\varepsilon_t\}$  is such that  $z_t \equiv \varepsilon_t/\sigma_t \sim IID(0,1)$ ; semi-strong GARCH holds where  $\{\varepsilon_t\}$  is such that  $E[\varepsilon_t|\varepsilon_{t-1},\ldots] = 0$  and  $E[\varepsilon_t^2|\varepsilon_{t-1},\ldots] = \sigma_t^2$ ; weak GARCH holds where  $\{\varepsilon_t\}$  is such that  $P[\varepsilon_t|\varepsilon_{t-1},\ldots] = 0$  and  $P[\varepsilon_t^2|\varepsilon_{t-1},\ldots] = \sigma_t^2$ , where  $P[\varepsilon_t^2|\varepsilon_{t-1},\ldots]$  denotes the best linear predictor of  $\varepsilon_t^2$  given a constant and past values of both  $\varepsilon_t$  and  $\varepsilon_t^2$ .

#### 2.2 Estimation by aggregation

Drost and Nijman (1993) showed that time-aggregated weak GARCH processes lead to processes of the same class, and gave deterministic relations between the coefficients (and the kurtosis) of the high frequency process and corresponding time-aggregated (lowfrequency) process for the weak GARCH (1,1) case. As Drost and Nijman noted, such relations can in principle be used to obtain estimates of the parameters at one frequency from those at another. In this section we examine the strategy of low-frequency estimation based on prior high-frequency estimates. Time aggregation relations of course differ for stock and flow variables; here we treat flows, such as asset returns.

Consider the high-frequency GARCH(1,1) process

$$\sigma_{\tau}^2 = \omega + \alpha_1 \varepsilon_{\tau-1}^2 + \beta_1 \sigma_{\tau-1}^2; \qquad (2.2.1)$$

if  $\varepsilon_{(h)t} = \sum_{j=t(h-1)+1}^{th} \varepsilon_j$  is the aggregated flow variable, then its volatility at the low frequency follows the weak GARCH(1,1) process

$$\sigma_{(h)t}^2 = \overline{\alpha}_0 + \overline{\alpha}_1 \varepsilon_{(h)t}^2 + \overline{\beta}_1 \sigma_{(h)t-1}^2, \qquad (2.2.2)$$

with  $\overline{\alpha}_0, \overline{\alpha}_1, \overline{\beta}_1$  given by the corresponding formulae (13-15) for  $\overline{\psi}, \overline{\alpha}, \overline{\beta}$  in Drost and Nijman (1993), adjusting for notation. To obtain consistent estimation by QML of the high-frequency model, it will be necessary that the process is semi-strong GARCH: the standard Quasi-ML estimator of the GARCH model will in general be inconsistent in weak GARCH models (as noted by Meddahi and Renault 1996, 2000 and Francq and Zakoïan 1998; see the latter reference for an example and M-R 2000 for a Monte Carlo example on samples of 80 000 – 150 000 simulated low frequency observations).

We will show that the mapping

$$\begin{pmatrix} \overline{\alpha}_0 \\ \overline{\alpha}_1 \\ \overline{\beta}_1 \end{pmatrix} = \psi \begin{pmatrix} \omega \\ \alpha_1 \\ \beta_1 \end{pmatrix}$$
(2.2.3)

provided by the Drost-Nijman formulae is a continuously differentiable mapping; it is also analytic over the region where the parameters are defined.

This implies that any consistent estimator of the high-frequency parameters  $(\omega, \alpha_1, \beta_1)$ leads to a consistent estimator of the low-frequency parameters  $(\overline{\alpha}_0, \overline{\alpha}_1, \overline{\beta}_1)$ , and similarly that an asymptotically Normal estimator of the high-frequency parameters results in asymptotic Normality of the low-frequency parameters.

Denote the vector 
$$\begin{pmatrix} \omega \\ \alpha_1 \\ \beta_1 \end{pmatrix}$$
 by  $\eta$  and, correspondingly, let  $\overline{\eta} = \begin{pmatrix} \overline{\alpha}_0 \\ \overline{\alpha}_1 \\ \overline{\beta}_1 \end{pmatrix}$ . Then  $\psi(\eta) = \overline{\eta}$ .

Now denote by  $\Omega \in \mathcal{R}^3$  the region

$$\Omega = \{ (\omega, \alpha_1, \beta_1) \in \mathcal{R}^3 | \ \omega > 0, \ \alpha_1 \ge 0, \ \beta_1 \ge 0; \ \alpha_1 + \beta_1 < 1 \},\$$

that is, the region for which the GARCH(1,1) process is defined (see, e.g., Bollerslev 1986).

Theorem 1. For any estimator  $\hat{\eta}$  of  $\eta$  such that (i)  $\hat{\eta} \xrightarrow{p} \eta$ ; (ii)  $\hat{\eta} \stackrel{a}{\sim} N(\eta, V(\eta))$ , the estimator  $\hat{\overline{\eta}} = \psi(\hat{\eta})$  is such that for  $\hat{\eta}$  satisfying (i),

 $\hat{\overline{\eta}} \xrightarrow{p} \overline{\eta},$ 

and for  $\hat{\eta}$  satisfying (ii),

$$\hat{\overline{\eta}} \stackrel{a}{\sim} N(\eta, V(\hat{\overline{\eta}})),$$

where the asymptotic covariance matrix is  $V(\hat{\eta}) = \frac{\partial \psi}{\partial \eta'} V(\eta) \frac{\partial \psi'}{\partial \eta}$ .

Proof. It follows from (i) and consequently also from (ii) that since  $\eta \in \Omega$ ,  $P(\hat{\eta} \in \Omega) \to 1$ . Consider now the formulae for  $\overline{\eta} = \psi(\eta)$  over  $\Omega$  in Drost and Nijman (1993). From (15) of D-N we can obtain  $\overline{\beta}_1$  from a solution to a quadratic equation of the form  $Z^2 - cZ + 1 = 0$ , where

$$c = c(\omega, \alpha_1, \beta_1, \kappa) \tag{2.2.4}$$

is obtained from the expression in (15) of D-N. For  $\eta \in \Omega$  it follows that c > 2 and therefore  $\overline{\beta}_1 = \frac{c}{2} - \left[ (\frac{c}{2})^2 - 1 \right]^{\frac{1}{2}}$  is such that  $0 < \overline{\beta}_1 < 1$ . Moreover, it can be shown that  $\overline{\beta}_1 < (\alpha_1 + \beta_1)^h$  and so  $\overline{\alpha}_1$  obtained from (13) in D-N also lies between 0 and 1.

The transformation  $\psi \begin{pmatrix} \omega \\ \alpha_1 \\ \beta_1 \end{pmatrix}$  can be written as

$$\psi\begin{pmatrix} \omega\\ \alpha_1\\ \beta_1 \end{pmatrix} = \begin{pmatrix} h\omega \frac{1-(\beta_1+\alpha_1)^h}{1-(\beta_1+\alpha_1)}\\ (\beta_1+\alpha_1)^h - \frac{c}{2} + \left[ (\frac{c}{2})^2 - 1 \right]^{\frac{1}{2}}\\ -\frac{c}{2} + \left[ (\frac{c}{2})^2 - 1 \right]^{\frac{1}{2}} \end{pmatrix},$$

where c is given by (2.2.4); it is defined and differentiable everywhere in  $\Omega$ .

Note that (as follows from Drost and Nijman 1993), even if  $\beta_1 = 0$ ,  $\overline{\beta}_1$  is nonzero as long as  $\alpha > 0$ . As *h* increases,  $\overline{\alpha}_1$  and  $\overline{\beta}_1$  decline; given  $\alpha_1$  and  $\beta_1$ , conditional heteroskedasticity vanishes for sufficiently large *h*. Therefore, for substantial conditional heteroskedasticity to be present in the low-frequency (aggregated) flow process,  $\alpha_1 + \beta_1$ must be close to unity. Suppose now that a standard quasi-Maximum Likelihood estimator is used with semistrong GARCH high-frequency data to obtain estimates of  $\eta$ . Its asymptotic covariance matrix is  $V[\hat{\eta}_{QML}] = [W'W]^{-1}B'B[W'W]^{-1}$ , where

$$W'W = \sum_{\tau=1}^{h} T\left[\frac{g_{\tau}}{\sigma_{\tau}^{2}}\right] \left[\frac{g_{\tau}}{\sigma_{\tau}^{2}}\right]'$$

and 
$$B'B = \sum_{\tau=1}^{T} \left[ \frac{\varepsilon_{\tau}^2}{\sigma_{\tau}^2} - 1 \right]^2 \left[ \frac{g_{\tau}}{\sigma_{\tau}^2} \right] \left[ \frac{g_{\tau}}{\sigma_{\tau}^2} \right]'$$

with  $g_{\tau} = \frac{\partial \sigma_{\tau}^2}{\partial \eta} = \begin{pmatrix} 1 \\ \varepsilon_{\tau-1}^2 \\ \sigma_{\tau-1}^2 \end{pmatrix}$ . The asymptotic variance of the estimator  $\hat{\eta}$  based on flow

aggregation is then

$$\frac{\partial \psi}{\partial \eta'} [W'W]^{-1} B' B [W'W]^{-1} \frac{\partial \psi'}{\partial \eta}.$$
(2.2.5)

If  $\hat{\eta}_{QML}$  is the MLE this reduces to

$$\frac{\partial \psi}{\partial \eta'} [W'W]^{-1} \frac{\partial \psi'}{\partial \eta}.$$
(2.2.6)

Example 1. Let the high-frequency process be ARCH(1); aggregation then leads to a weak GARCH(1,1) process for the low-frequency data. However, the asymptotic covariance matrix for the estimator  $\hat{\eta}$  is of rank 2 rather than 3, since the middle part in (2.2.5) or (2.2.6) is of dimension 2 × 2. This indicates that there are cases where  $\hat{\eta}$  is clearly more efficient than  $\bar{\eta}_{ML}$  (or  $\bar{\eta}_{QML}$ ) based on low-frequency data alone, with covariance matrix of rank 3.

While estimation is feasible by this method, the requirements of this strategy, even for consistent estimation, are fairly severe. In particular, the potential inconsistency of QML estimation when only weak GARCH conditions apply means that we must assume semi-strong GARCH at the high frequency if estimation is by QML. This is, however, an arbitrary specification; if the high frequency data are themselves aggregates of yet higher frequencies, the semi-strong conditions do not follow. While consistent estimation of weak GARCH models is in principle possible (see Francq and Zakoïan 1998), the QML estimator does not accomplish this.

More generally, estimation based on aggregation presumes knowledge of the highfrequency structure, and requires the computation of different aggregation formulae for each model form to be estimated. For these reasons, we will proceed to investigate estimators based on the integrated volatility, which do not presume knowledge of the highfrequency process beyond the conditions necessary for consistency of the daily volatility estimate.

#### 2.3 Estimation by regression using integrated volatility

As noted above, models of the form (2.1.1) and (2.1.2) are directly estimable if we have estimates of the conditional variance of the low-frequency observations,  $\sigma_t^2$ , for example from the daily integrated volatility, as in Andersen and Bollerslev.<sup>1</sup> However, we will not follow Andersen and Bollerslev in treating the observation as exact. Instead, we introduce into the model the measurement error arising in estimation of  $\overline{\sigma}_t^2$  from the daily integrated volatility (2.1.0), a specification also employed by Maheu and McCurdy (2000). Let

$$\hat{\sigma}_t^2 = \sigma_t^2 + e_t; \tag{2.3.1}$$

properties of  $\{e_t\}$  follow from (2.1.0) and will be considered below.

The ARCH and GARCH models become

$$\hat{\sigma}_t^2 = \omega + \sum_{i=1}^q \alpha_i \varepsilon_{t-i}^2 + e_t, \qquad (2.3.2)$$

$$\hat{\sigma}_t^2 = \omega + \sum_{i=1}^q \alpha_i \varepsilon_{t-i}^2 + \sum_{i=1}^p \beta_i \hat{\sigma}_{t-i}^2 - \sum_{i=1}^p \beta_i e_{t-i} + e_t.$$
(2.3.3)

Both (2.3.2) and (2.3.3) are in principle estimable as regression models. The model (2.3.3) has an error term with an MA(s) form; the coefficients of this moving average process are subject to the constraint embodied in (2.3.3) that they are the same (up to sign) as the coefficients on lagged values of  $\hat{\sigma}_t^2$ . Estimation of these models by LS or QML does however require relatively strong moment conditions to hold on the regressors and the volatility estimation errors  $\{e_t\}$ ; note that this is unlike the standard GARCH model estimated by QML where conditions are usually applied to the rescaled squared innovations.

Maheu and McCurdy (2000) find good results using the constrained model, estimated by QML, on foreign exchange returns. Bollen and Inder (1998) estimate a model similar to (2.3.3) by standard QML methods (without accounting for the error autocorrelation structure), using intra-day data to obtain estimates of an unobservable sequence related to daily volatility. This approach requires consistency of the estimates of the unobservable

<sup>&</sup>lt;sup>1</sup>Since we will be discussing low-frequency parameters hereafter, we no longer need to distinguish from high-frequency and will omit the 'bar' in symbols, referring to  $\sigma^2$ ,  $\alpha$ ,  $\beta$ , etc. for the low-frequency values.

sequence as the number of intra-day observations per day increases without bound, to obtain consistency of the estimator; Bollen and Inder find good results on a sample of S & P Index futures with a large number of observations per day.

Here we will consider estimation of models having the ARCH model (2.3.2), followed by computation of GARCH parameters from the ARCH approximation. The method that we will use should produce estimates which are relatively insensitive to the error distribution and to existence of moments of the errors, which is particularly advantageous in financial data. This strategy also has the advantage of producing immediately an estimated model which is directly useable for forecasting, and of allowing computation of parameters of any GARCH(p, q) model from a given estimated ARCH representation. Sufficient conditions for consistent and asymptotically normal estimation are given in Section 3 below; it is not necessary that the number of intra-day observations per day (h) increase without bound.

To obtain estimates of GARCH parameters from the ARCH representation we pursue an estimation strategy related to that of Galbraith and Zinde-Walsh (1994, 1997), in which autoregressive models are used in estimation of MA or ARMA models. Here, a high-order ARCH model is used, and estimates of GARCH (p,q) parameters are deduced from the patterns of ARCH coefficients. This is another example of what Galbraith and Zinde-Walsh (2001) refer to as 'analytical indirect inference', in that a method analogous to indirect inference (Gouriéroux et al. 1993) is used, but where the mapping from estimates of the auxiliary model to the model to be estimated is obtained analytically, rather than by simulation.

Consider first a case of known conditional variance  $\sigma_t$ . The GARCH process (2.1.2) has a form analogous to the ARMA(p,q); using standard results on representation of an ARMA (p,q) process in MA form (see, e.g., Fuller 1976), we can express (2.1.2) in the form

$$\sigma_t^2 = \kappa + \sum_{\ell=1}^{\infty} \nu_\ell \varepsilon_{t-\ell}^2, \qquad (2.3.4)$$

with  $\nu_0 = 0$  and

$$\nu_{1} = \alpha_{1}$$

$$\nu_{2} = \alpha_{2} + \beta_{1}\nu_{1}$$

$$\vdots$$

$$\nu_{\ell} = \alpha_{\ell} + \sum_{i=1}^{\min(\ell,p)} \beta_{i}\nu_{\ell-i}, \quad \ell \leq q,$$

$$\nu_{\ell} = \sum_{i=1}^{\min(\ell,p)} \beta_{i}\nu_{\ell-i}, \quad \ell > q;$$
(2.3.5)

and finally

$$\kappa = (1 - \beta(1))^{-1}\omega = \left(1 - \sum_{i=1}^{p} \beta_i\right)^{-1}\omega.$$
 (2.3.6)

Giraitis et al. (2000) give general conditions under which the  $ARCH(\infty)$  representation is possible for the strong GARCH(p,q) case; only the existence of the first moment and summability of the coefficients  $\nu_{\ell}$  (in our notation) are required for the existence of a strictly stationary  $ARCH(\infty)$  solution as given in (2.3.4).

To estimate the model using a truncated version of this  $ARCH(\infty)$  representation, we use the estimated low-frequency conditional variance from (2.1.0), defining the estimation error as in (2.3.1) and substituting into (2.3.4) to obtain

$$\hat{\sigma}_{t}^{2} = \kappa + \sum_{\ell=1}^{k} \nu_{\ell} \varepsilon_{t-\ell}^{2} + e_{t}.$$
(2.3.7)

The truncation parameter k must be such that  $k \to \infty$ ,  $k/T \to 0$  for consistent estimation of the GARCH model.

The model (2.3.7) may be estimated by LS or, to obtain results robust to less restrictive conditions on the volatility estimation errors, LAD. Asymptotic properties of the LAD estimator are considered in Section 3. Estimation proceeds by first obtaining estimates of  $\beta = (\beta_1, \beta_2, \ldots, \beta_p)$  from (2.3.5) for  $\ell > q$ , followed by estimation of the q parameters of  $\alpha$ from the first q relations of (2.3.5), and of  $\omega$  from (2.3.6).

Begin by defining

$$v(0) = \begin{bmatrix} \nu_{q+1} \\ \nu_{q+2} \\ \vdots \\ \nu_k \end{bmatrix}, \text{ and } v(-i) = \begin{bmatrix} \nu_{q+1-i} \\ \nu_{q+2-i} \\ \vdots \\ \nu_{k-i} \end{bmatrix}.$$
 (2.3.8)

Next define the  $(k-q)\times p$  matrix  $V=[v(-1)v(-2)\ldots v(-p)]=$ 

$\nu_q$	$\nu_{q-1}$		$\nu_{q-p+1}$
$\nu_{q+1}$	$ u_q$		$\nu_{q-p+2}$
			:
$\lfloor \nu_{k-1}$	$\nu_{k-2}$		$\begin{bmatrix} \cdot \\ \nu_{k-p} \end{bmatrix}$
$ \nu_{\kappa-1} $	$\nu_{\kappa-2}$	• • •	$\nu_{\kappa}-p$ -

where  $\nu_r = 0$  for  $r \leq 0$ . It follows from (2.3.5) that  $v(0) = \beta' V$ .

The  $p \times 1$  vector of estimates  $\hat{\beta}$  is defined by

$$\hat{\beta} = (\hat{V}'\hat{V})^{-1}\hat{V}'\hat{v}(0), \qquad (2.3.9)$$

where the circumflex indicates replacement of  $\nu_{\ell}$  with the OLS-estimated values  $\hat{\nu}_{\ell}$  in the definitions above. An estimate of  $\alpha$  can then be obtained using the estimate of  $\beta$  and the relations (2.3.5): that is

$$\hat{\alpha}_{1} = \hat{\nu}_{1} 
\hat{\alpha}_{2} = \hat{\nu}_{2} - \hat{\beta}_{1}\hat{\nu}_{1} 
\vdots 
\hat{\alpha}_{q} = \hat{\nu}_{q} - \sum_{i=1}^{\min(q,p)} \hat{\beta}_{i}\hat{\nu}_{q-i}.$$
(2.3.10)

Finally,

$$\hat{\omega} = \hat{\kappa}(1 - \hat{\beta}(1)) = \hat{\kappa} \left(1 - \sum_{i=1}^{p} \hat{\beta}_i\right).$$

The covariance matrix of the estimates can be obtained easily from the estimates of the representation (2.3.7) and the Jacobian of the transformation. Let the parameter vector be  $\delta = (\omega, \beta, \alpha)$ , and let  $\psi^2$  be the variance of the noise  $e_t$  in (2.3.1). Then

$$\operatorname{var}(\delta) = J'(\operatorname{var}\hat{\nu})J,$$

where  $\hat{\nu}$  is the vector of estimated ARCH parameters and J is the Jacobian of the transformation (2.3.9)-(2.3.10). Computation of the covariance matrix of the LAD parameter vector is discussed in section 3.

#### 3. Asymptotic properties of the integrated volatility-regression estimates

In this section we discuss conditions for consistent and asymptotically Normal estimation of the integrated volatility-regression model of Section 2 by LAD. Results for QML estimation of the ARCH model were established by Weiss (1986), using the assumption of finite fourth moments of the unnormalized data. Lumsdaine (1996) established consistency and asymptotic Normality of the QMLE for GARCH models by imposing conditions on the re-scaled data,  $z_t = \varepsilon_t/\sigma_t$ , including the IID assumption and the existence of high-order moments. Lee and Hansen (1994) generalized these results to  $\{z_t\}$  which are not IID, but simply strictly stationary and ergodic.

Recall that the ARCH(k)-type model (2.3.7) is a regression model, and consider for each T the  $\sigma$ -field  $F_t$  generated by the regressors  $\{(\varepsilon_t^2, \varepsilon_{t-1}^2, \ldots, \varepsilon_{t-k}^2), t = k+1, \ldots, T.\}$ . Denote by  $W_T$  the symmetric matrix with elements  $w_{ij} = T^{-c}(\sum_{t=1}^T \varepsilon_{t-i}^2 \varepsilon_{t-j}^2)$ , and let  $V_T = W_T^{\frac{1}{2}}$ .

Assumption 1. There exists some c such that

$$w_{ij} = O_p(1); \quad W_T^{-1} = O_p(1); \quad (T^{-\frac{c}{2}}) \max \varepsilon_t^2 = o_p(1).$$

The assumption is trivially satisfied with c = 1 if  $\varepsilon^2$  possesses a sufficient number of moments, but the assumption does not require the existence of moments.<sup>2</sup>

Next consider an assumption on the daily volatility estimation error,

$$e_t = \hat{\sigma}_t^2 - \sigma_t^2. \tag{3.1.1}$$

The following assumption embodies both the Error Assumption of Pollard (1991, p.189) and the additional assumption of Pollard's Theorem 2 that the realizations of the process and the errors (here, volatility estimation errors) are assumed independent.

Assumption 2. The volatility estimation errors  $\{e_t\}$  are IID with median 0 and a continuous positive density f(.) in the neighbourhood of zero. The sequences of errors  $\{e_t\}$  and of realizations of the process  $\{\varepsilon_t\}$  are independent.

Theorem 2. Consider the model (2.3.7),

$$\hat{\sigma}_t^2 = \kappa + \sum_{\ell=1}^k \nu_\ell \varepsilon_{t-\ell}^2 + e_t$$

and suppose that Assumptions 1 and 2 are satisfied. Then if  $\hat{\nu} = (\hat{\kappa}, \hat{\nu}_1, \dots, \hat{\nu}_k)$  is the the LAD-estimated parameter vector,

$$2f(0)V_T(\hat{\nu}-\nu) \xrightarrow{D} N(0, I_{k+1}).$$

**Proof.** Follows from Pollard (1991, Theorem 2 and Example 1) for c = 1; can be extended to any c. Assumption 1 satisfies the conditions (ii)-(iv) of Theorem 2 of Pollard (1991), and combined with Assumption 2 provides all of the conditions (i)-(iv) for the asymptotic distribution to hold.

#### 4. Simulation evidence

In this section we present evidence primarily on the finite-sample performance of the regression estimator of Section 2.3 using the daily integrated volatility, and for comparison the standard Quasi-ML estimator based on daily data alone. The QML procedure

<sup>&</sup>lt;sup>2</sup>As an example of the former point, consider a case where the eighth unconditional moment of  $\varepsilon_t$  exists. Then (i) is satisfied for c = 1 by the WLLN,  $T^{-1}(\sum_{t=1}^T \varepsilon_{t-i}^2 \varepsilon_{t-j}^2) \xrightarrow{p} E(\varepsilon_{t-i}^2 \varepsilon_{t-j}^2)$ . At the same time, if we re-write the ARCH(k) model (2.3.7) as a stationary AR(k) model by defining  $w_t = \varepsilon_t^2 - \sigma_t^2$  (note that  $E(w_t | \varepsilon_t^2, \varepsilon_{t-1}^2, \ldots) = 0$  and  $\operatorname{var}(w_t) < \infty$ ), we obtain  $\varepsilon_t^2 = \kappa + \sum_{\ell=1}^k \nu_\ell \varepsilon_{t-\ell}^2 + w_t$ . Following example 2 of Pollard (1991) (generalizing to AR(k)), it follows that  $\max \varepsilon_t^2 = o_p(T^{1/2})$ . Of course, Assumption 1 can also hold in cases where moments do not exist.

described and implemented by Aptech Systems (1998) is used for the standard estimates. The regression estimator uses (2.1.0) for a daily volatility estimate, followed by estimation of (2.3.7) by OLS or LAD, and transformation to GARCH parameter estimates via (2.3.9)-(2.3.10).

In the first set of simulation exercises, low-frequency (daily) data alone are simulated. The innovations are Normal, leading to a strong GARCH model estimated by a true ML estimator. These cases are therefore as favourable as possible for the QML estimator. The low-frequency sample size T is set at {200,600} and the number of replications is 2000 for each experiment. Although the first set of experiments uses simulated low-frequency data only, the parameter values are chosen for compatibility with the later experiments; the four sets of parameter values used result from the aggregation of high-frequency GARCH processes having parameters { $\omega, \alpha, \beta$ } equal to (.01, .018, .98), (.01, .05, .945), (.01, .08, .89) and (.01, .10, .85). The corresponding low-frequency parameters are computed from the aggregation formula of Drost and Nijman (1993) for the GARCH (1,1) flow case; the values are (6.102, .0555, .8957), (5.889, .1371, .7450), (4.442, .0736, .3934), and (3.613, .0540, .2234).

In a second set of simulations, for the same sample sizes and number of replications, the high-frequency GARCH process is simulated directly as in the first experiments, (strong GARCH, with normal errors) and aggregated to form a (weak GARCH) low-frequency (daily) returns process. Estimates of the daily GARCH parameters on these daily data are obtained by QML and the regression estimator.

As just indicated, the 'true' daily parameters are computed from the aggregation formula of Drost and Nijman (1993), for comparison with the estimates from each method. Note that QML applied directly to the daily data is now technically inconsistent because only the weak GARCH conditions can be guaranteed to apply. For the regression estimates, the a daily volatility estimate has a noise variance implied by the number of observations per day, h, which in these examples is set at 25 to keep overall sample sizes manageable. This number, while modest, is compatible with the suggestions of the more recent literature on integrated volatility (Andersen et al. 1999) which suggests that the optimal frequency to use in estimation of daily volatility is not necessarily the highest frequency available. A small value, representing a relatively noisy estimate of daily volatility, is in general less favourable for the relative performance of the regression estimator.

Because the QML and LAD-ARCH estimators are evaluated here using different information sets—the LAD-ARCH uses the higher-frequency data, while QML does not the outcome of any comparison depends upon the information content of the additional (intra-day) data. The comparison here is made on parameter values that appear to be representative of empirical outcomes, and uses a relatively modest information content of intra-day data. Nonetheless, such a comparison can only be viewed as illustrative. In experiments of the type performed here, using Normal innovations, QML or LAD-ARCH can dominate the other in a finite sample as the estimates of realized volatility become (respectively) relatively noisy or relatively precise. Of course, very heavy-tailed or skewed error distributions will tend further to favour the LAD-ARCH estimator.

Results from the first set of simulations are contained in Figure 1a/b Results for the second set of simulations are reported in Figure 2a/b and 3a/b. With the second set of simulations we also report a forecast comparison of 1-step-ahead out-of-sample conditional volatility forecasts from each estimation method, evaluated by RMS forecast error.

Begin by comparing corresponding quadrants of Figure 1a with those of Figure 2a, and of 1b with 2b. The QML estimator's performance is little different between these two cases, despite the fact that the Figure 1 cases are in fact true ML estimators, and that the Figure 2 cases show QML in a weak GARCH case where the estimator is not in general consistent. However these results, in line with those mentioned by Drost and Nijman (1993, p. 922), suggest that the QMLE may be converging to values quite close to the true parameter values.

Nonetheless the finite-sample performance of the QML estimator in these cases shows problems which are of a fairly standard type, not specific to GARCH estimation. The QMLE's concentration of probability mass near the boundaries of the identification region are typical of the 'pile-up' problems familiar from cases such as ML estimation of MA parameters; here the QMLE performs poorly when  $\alpha$  is near zero. Estimates of  $\beta$  are particularly poor in these examples, showing large probability masses in the vicinity of 0.9 regardless of the true value of  $\beta$ , in each of the three cases where  $\alpha < 0.1$ . Where  $\alpha = 0.1373$ , by contrast, estimates of  $\beta$  are concentrated near the true value, particularly at the larger sample size. Of course, QMLE performance improves in general for larger values of  $\alpha$ , as we move away from the boundary of the identification region; the values chosen here, however, are typical of those appearing in the empirical literature.

Figure 3 contains results for the LAD-ARCH estimator in corresponding cases, for h = 25: that is, using the relatively low high-frequency information content noted above. Nonetheless, additional information is being exploited, and we might expect that the estimator should perform relatively well. This expectation is borne out in Figure 3a, where LAD-ARCH estimates of  $\alpha$  show much better conformity with the Normal and smaller dispersion, as well as less pile-up near zero.

In Figure 3b, we see less dramatic gains in estimation of  $\beta$ ; conformity with the Normal is improved realtive to corresponding Figure 2a cases, but there remains substantial pile-up at  $\beta = 0$  and again we see dispersion of the estimates across much of the [0, 1] interval. However, the large spurious peak in the density near 0.9 is eliminated. Conformity with the Normal is markedly improved at T = 600. These cases, embodying Normal errors in the simulation DGP, do not offer the most favourable circumstances for comparison of LAD-ARCH; with substantially skewed error distributions, we would expect to see the traditional advantage of LAD estimation in robustness to skewness and heavy-tailed errors. However, the estimator does offer clear gains nonetheless.

#### 5. Concluding remarks

Estimation of GARCH models via integrated volatility is feasible even with a modest number of intra-day observations. Such estimates have a number of apparent advantages over those obtained from standard QML estimates, which can be quite erratic in small samples. The use of integrated volatility in this way also opens up other avenues for estimation, including estimation of quantiles of volatility other than the median.<sup>3</sup>

 $<sup>^3{\</sup>rm This}$  has been investigated for ARCH models estimated with standard (single-frequency) data by Koenker and Zhao (1996).

### References

Andersen, T.G. and T. Bollerslev (1997) Intraday Periodicity and Volatility Persistence in Financial Markets. *Journal of Empirical Finance* 4, 115-158.

Andersen, T.G. and T. Bollerslev (1998) Answering the Skeptics: Yes, Standard Volatility Models Do Provide Accurate Forecasts. International Economic Review 39(4), 885-905.

Andersen, T.G., T. Bollerslev, F.X.Diebold and P.Labys (1999) The Distribution of Exchange Rate Volatility. Working paper.

Aptech Systems (1998) FANPAC. Maple Valley, WA.

Baillie, R.T., T. Bollerslev, and H.O. Mikkelsen (1996) Fractionally Integrated Generalized Autoregressive Conditional Heteroskedasticity. *Journal of Econometrics* 74, 3-30.

Barndorff-Neilsen, O.E. and N. Shephard (2000) Econometric analysis of realised volatility and its use in estimating Lévy-based non-Gaussian OU type stochastic volatility models. Working paper, Nuffield College, Oxford.

Bollen, B. and B. Inder (1998) A General Volatility Framework and the Generalised Historical Volatility Estimator. Working paper, Monash University.

Bollerslev, T. (1986) Generalized Autoregressive Conditional Heteroskedasticity. Journal of Econometrics 31, 307-327.

Drost, F.C. and B.J.M. Werker (1996) Closing the GARCH Gap: Continuous Time GARCH Modeling. *Journal of Econometrics* 74, 31-57.

Drost, F.C. and T.E. Nijman (1993) Temporal Aggregation of GARCH Processes. Econometrica 61(4), 909-927.

Engle, R.F. (1982) Autoregressive Conditional Heteroskedasticity with Estimates of the Variance of U.K. Inflation. *Econometrica* 50, 987-1008

Francq, C. and J.-M. Zakoïan (1998) Estimating Weak GARCH Representations. Working paper, Université du Littoral–Cte d'Opale, Université de Lille I.

Fuller, W.A. (1976) Introduction to Statistical Time Series. Wiley, New York.

Galbraith, J.W. and V. Zinde-Walsh (1994) A Simple, Non-iterative Estimator for Moving-Average Models. *Biometrika* 81, 143–155.

Galbraith, J.W. and V. Zinde-Walsh (1997) On Some Simple, Autoregression-based Estimation and Identification Techniques for ARMA Models. *Biometrika* 84, 685–696.

Galbraith, J.W. and V. Zinde-Walsh (2001) Analytical Indirect Inference. Working paper, McGill University.

Giraitis, L., P. Kokoszka and R. Leipus (2000) Stationary ARCH Models: Dependence Structure and Central Limit Theorem. *Econometric Theory* 16, 3-22.

Gouriéroux, C. and A. Montfort (1993) Indirect Inference. Journal of Applied Econometrics 8, 85-118.

He, C. and T. Teräsvirta (1999) Fourth Moment Structure of the GARCH(p,q) process. *Econometric Theory* 15, 824-846.

Koenker, R. and Q. Zhao (1996) Conditional Quantile Estimation and Inference for ARCH Models. *Econometric Theory* 12, 793–813.

Lee, S.-W. and B.E. Hansen (1994) Asymptotic Theory for the GARCH(1,1) Quasi-Maximum Likelihood Estimator. *Econometric Theory* 10, 29-52.

Lumsdaine, R.L. (1996) Consistency and Asymptotic Normality of the Quasi-Maximum Likelihood Estimator in IGARCH(1,1) and Covariance Stationary GARCH(1,1) Models. Econometrica 64, 575-596.

Maheu, J.M. and T.H. McCurdy (2000) Nonlinear Features of Realized FX Volatility. Working paper, University of Alberta/ University of Toronto.

Meddahi, N. and E. Renault (1996) Aggregation and Marginalization of GARCH and Stochastic Volatility Models. Working paper 3597, CRDE, Université de Montréal.

Meddahi, N. and E. Renault (2000) Temporal Aggregation of Volatility Models. Working paper, Université de Montréal.

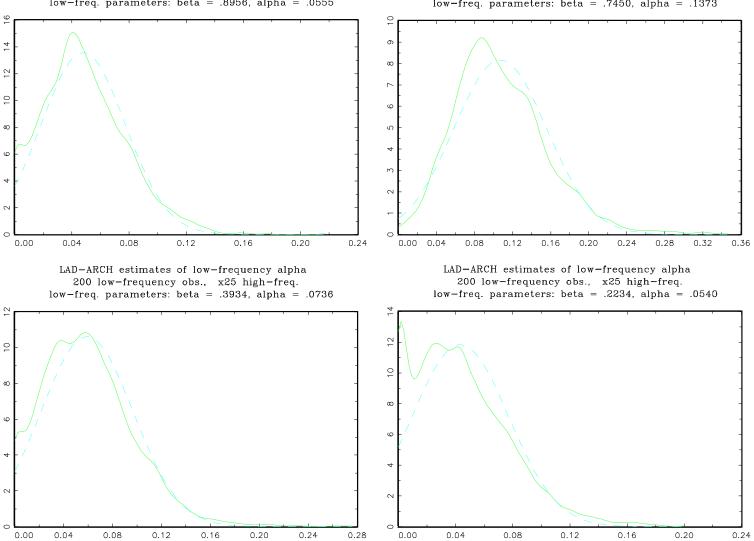
Nelson, D. (1992) Filtering and Forecasting with Misspecified ARCH Models I: Getting the Right Variance with the Wrong Model. *Journal of Econometrics* 52, 61-90.

Nelson, D. and D.P Foster (1995) Filtering and Forecasting with Misspecified ARCH models II: Making the Right Forecast with the Wrong Model. *Journal of Econometrics* 67, 303-335.

Phillips, P.C.B. (1995) Robust Nonstationary Regression. Econometric Theory 11, 912-951.

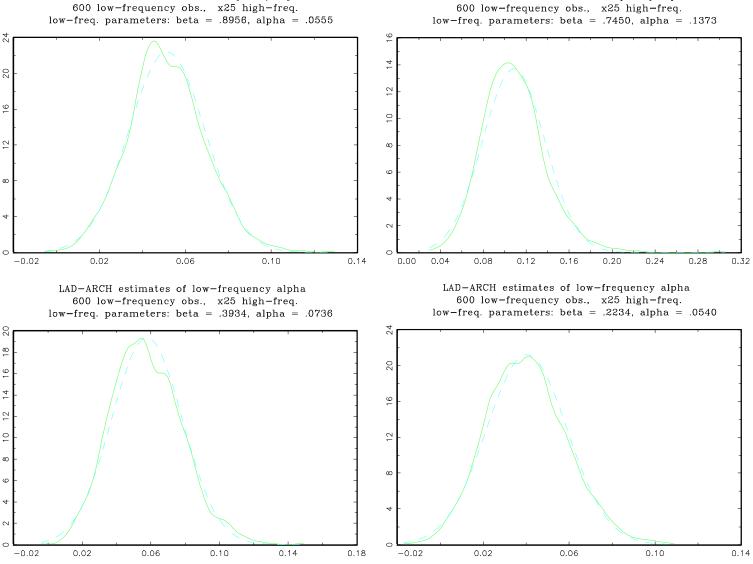
Pollard, D. (1991) Asymptotics for Least Absolute Deviation Regression Estimators. Econometric Theory 7, 186-199.

Weiss, A.A. (1986) Asymptotic Theory for ARCH Models: Estimation and Testing. Econometric Theory 2, 107–131.



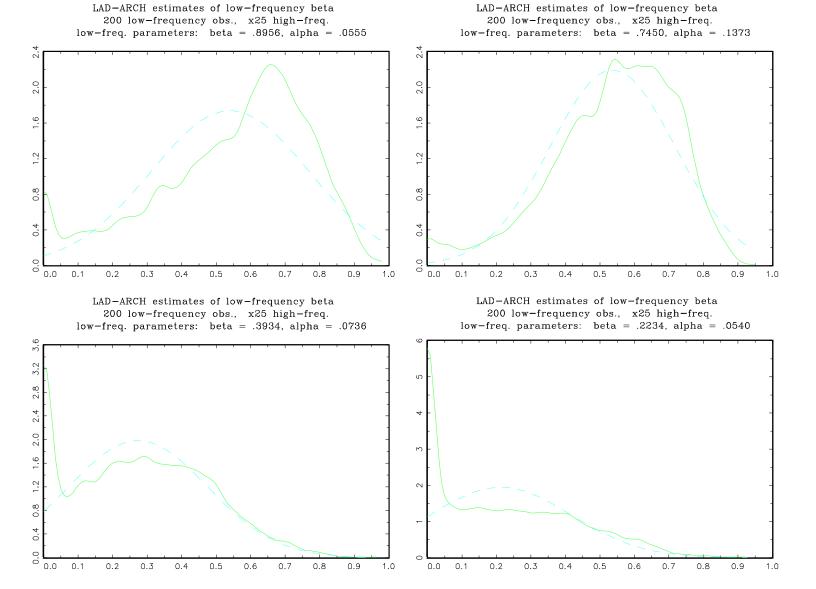
LAD-ARCH estimates of low-frequency alpha 200 low-frequency obs., x25 high-freq. low-freq. parameters: beta = .8956, alpha = .0555

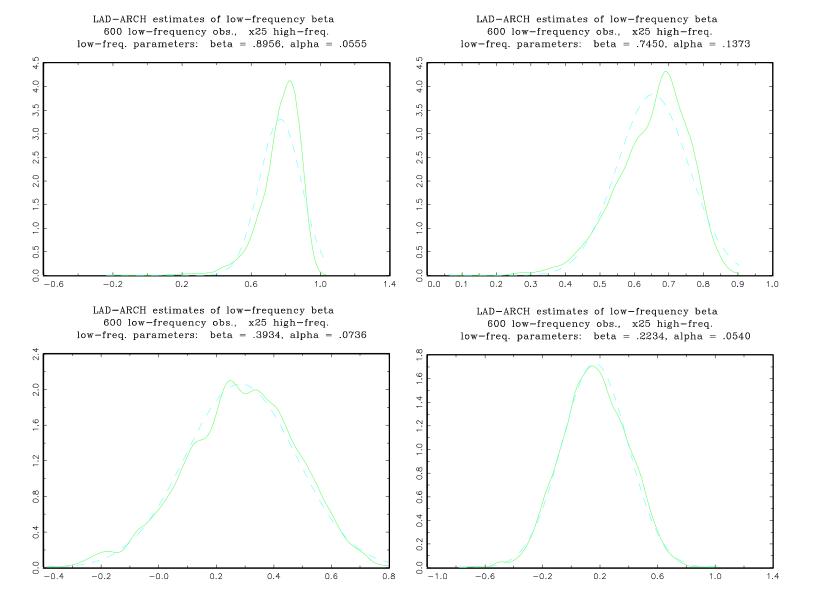
#### LAD-ARCH estimates of low-frequency alpha 200 low-frequency obs., x25 high-freq. low-freq. parameters: beta = .7450, alpha = .1373

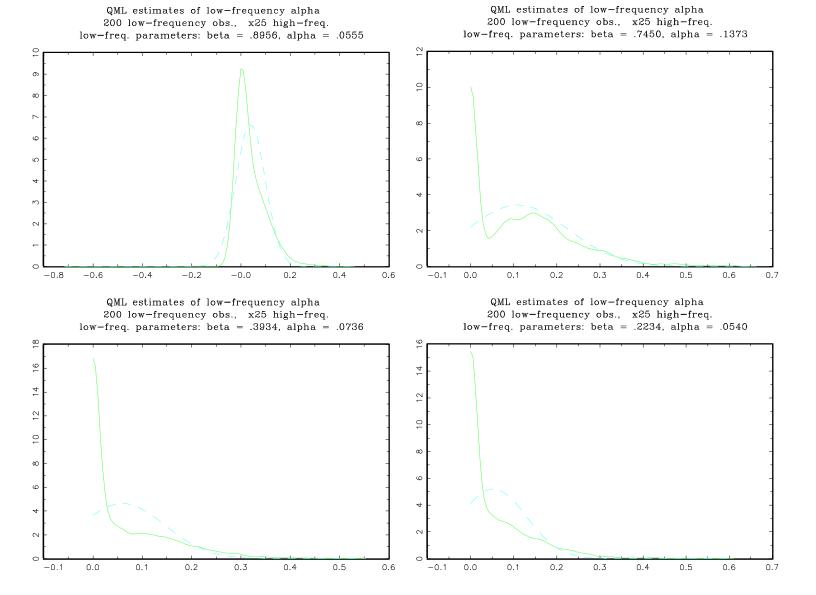


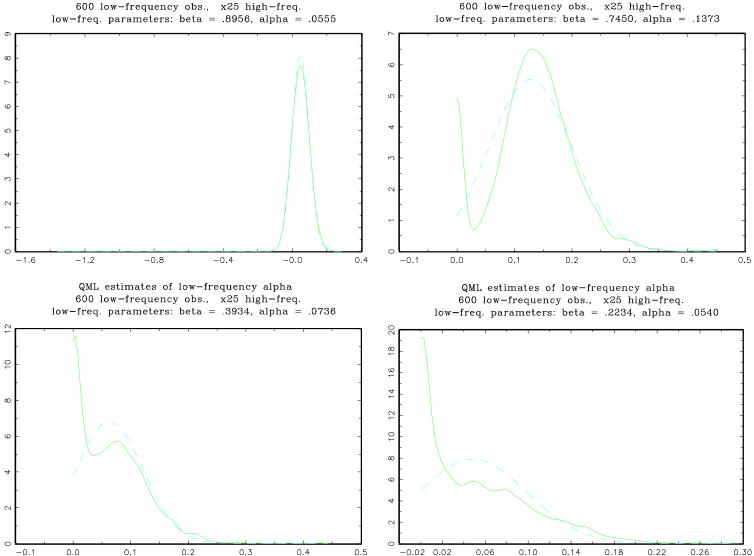
LAD-ARCH estimates of low-frequency alpha

LAD-ARCH estimates of low-frequency alpha 600 low-frequency obs., x25 high-freq. low-freq. parameters: beta = .8956, alpha = .0555



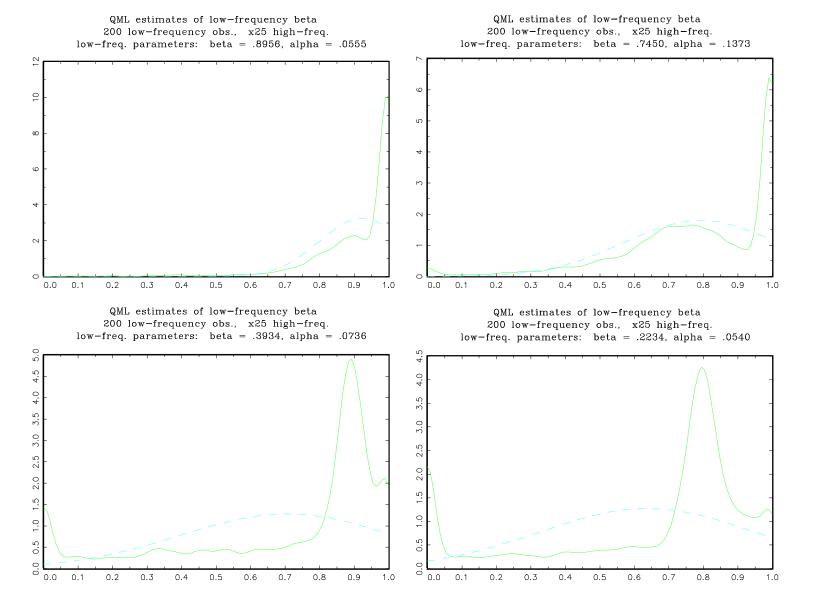


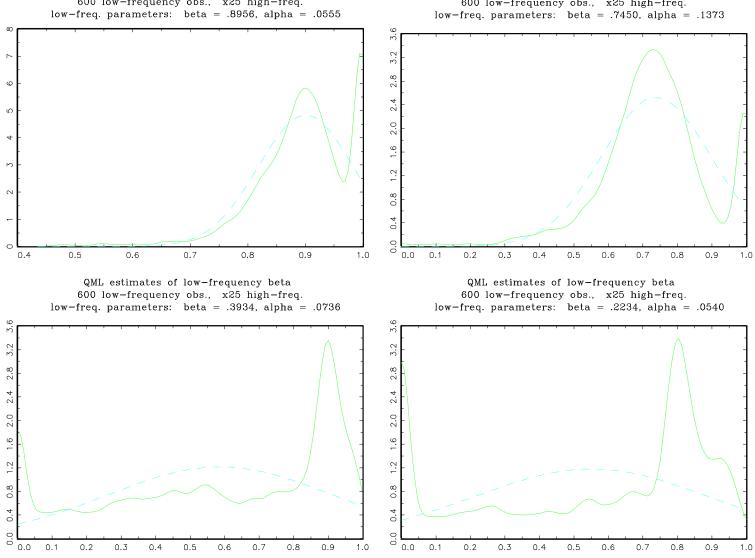




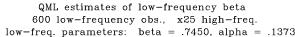
QML estimates of low-frequency alpha 600 low-frequency obs., x25 high-freq.

QML estimates of low-frequency alpha 600 low-frequency obs., x25 high-freq.





QML estimates of low-frequency beta 600 low-frequency obs., x25 high-freq.



# Liste des publications au CIRANO \*

# Cahiers CIRANO / CIRANO Papers (ISSN 1198-8169)

- 99c-1 Les Expos, l'OSM, les universités, les hôpitaux : Le coût d'un déficit de 400 000 emplois au Québec — Expos, Montréal Symphony Orchestra, Universities, Hospitals: The Cost of a 400,000-Job Shortfall in Québec / Marcel Boyer
- 96c-1 Peut-on créer des emplois en réglementant le temps de travail? / Robert Lacroix
- 95c-2 Anomalies de marché et sélection des titres au Canada / Richard Guay, Jean-François L'Her et Jean-Marc Suret
- 95c-1 La réglementation incitative / Marcel Boyer
- 94c-3 L'importance relative des gouvernements : causes, conséquences et organisations alternative / Claude Montmarquette
- 94c-2 Commercial Bankruptcy and Financial Reorganization in Canada / Jocelyn Martel
- 94c-1 Faire ou faire faire : La perspective de l'économie des organisations / Michel Patry

## Série Scientifique / Scientific Series (ISSN 1198-8177)

2001s-14	A Ricardian Model of the Tragedy of the Commons / Pierre Lasserre et Antoine
	Soubeyran
2001s-13	Carbon Credits for Forests and Forest Products / Robert D. Cairns et Pierre
	Lasserre
2001s-12	Estimating Nonseparable Preference Specifications for Asset Market Participants /
	Kris Jacobs
2001s-11	Autoregression-Based Estimators for ARFIMA Models / John Galbraith et
	Victoria Zinde-Walsh
2001s-10	Heterogeneous Returns to Human Capital and Dynamic Self-Selection / Christian
	Belzil et Jörgen Hansen
2001s-09	Return to a High School Diploma and the Decision to Drop Out: New Evidence
	from Canada / Daniel Parent
2001s-08	Leader and Follower: A Differential Game Model / Hassan Benchekroun et Ngo
	Van Long
2001s-07	Emission Taxes and Standards for an Asymmetric Oligopoly / Ngo Van Long et
	Antoine Soubeyran
2001s-06	Risque de modèle de volatilité / Ali Alami et Éric Renault
2001s-05	The Effect of Pay-for-Performance Contracts on Wages /Daniel Parent
2001s-04	Incentive Pay in the United States: Its Determinants and Its Effects / Daniel Parent

<sup>\*</sup> Vous pouvez consulter la liste complète des publications du CIRANO et les publications elles-mêmes sur notre site Internet à l'adresse suivante :