CREATES

Center for Research in Econometric Analysis of Time Series

CREATES Research Paper 2008-35

Bias-reduced estimation of long memory stochastic volatility

Per Frederiksen and Morten Ørregaard Nielsen

SOLIDUM PETIT IN PROFUNDIS · UNIVERSITAS ARHUSIENSIS ·

School of Economics and Management
University of Aarhus
Building 1322, DK-8000 Aarhus C
Denmark

Aarhus School of Business
University of Aarhus
Handelshøjskolen
Aarhus Universitet

UNIVERSITY OF COPENHAGEN

Bias-reduced estimation of long memory stochastic volatility*

Per Frederiksen†
Equity Trading & Derivatives, Nordea Markets

Morten Ørregaard Nielsen
Cornell University and CREATES

December 11, 2007

Abstract

We propose to use a variant of the local polynomial Whittle estimator to estimate the memory parameter in volatility for long memory stochastic volatility models with potential nonstationarity in the volatility process. We show that the estimator is asymptotically normal and capable of obtaining bias reduction as well as a rate of convergence arbitrarily close to the parametric rate, $n^{1/2}$. A Monte Carlo study is conducted to support the theoretical results, and an analysis of daily exchange rates demonstrates the empirical usefulness of the estimators.
JEL Classifications: C14, C22.
Keywords: Bias reduction, local Whittle estimation, long memory stochastic volatility model.

1 Introduction

The past two decades have witnessed increasing interest in fractionally integrated processes as a convenient way to model long memory properties of many time series. There is now a broad range of applications in e.g. finance and macroeconomics, see Baillie (1996) or Henry & Zaffaroni (2003) for some examples, and especially long memory in volatility has received considerable recent interest[1].

A popular way of modeling the observed persistence in volatility of financial returns is the long memory stochastic volatility (LMSV) model introduced by Breidt, Crato & de Lima (1998) and Harvey (1998). The LMSV model for financial returns takes the form

$$r_t = \kappa \exp\left(y_t/2\right) u_t, \tag{1}$$

*We are grateful to Torben G. Andersen, Niels Haldrup, Esben Høg, Asger Lunde, Frank S. Nielsen, two anonymous referees, an anonymous associate editor, and the co-editor Eric Ghysels for comments. This work was partly done while Frederiksen was visiting Northwestern University and Nielsen was visiting Queen's University and the University of Aarhus, their hospitality is gratefully acknowledged. We are grateful for financial support from the Danish Social Sciences Research Council (grant no. FSE 275-05-0220) and the Center for Econometric Analysis of Time Series (CREATES, funded by the Danish National Research Foundation).

†Please address correspondence to: Per Frederiksen, Equity Trading & Derivatives, Nordea Markets, 1401 Copenhagen C, Denmark; phone: +45 3333 6683; e-mail: per.frederiksen@nordea.com

[1] See, for example, Ding, Granger & Engle (1993), Baillie, Bollerslev & Mikkelsen (1996), Comte & Renault (1998), Ray & Tsay (2000), Andersen, Bollerslev, Diebold & Ebens (2001), Andersen, Bollerslev, Diebold & Labys (2001, 2003), Hurvich & Ray (2003), and Arteche (2004), among others.

where $\kappa > 0$ is a scale parameter, u_t is a white noise return shock with mean zero and unit variance, and y_t is a zero-mean, stationary log-volatility process with spectral density

$$f_y(\lambda) = \lambda^{-2d} \phi(\lambda), \tag{2}$$

where $\phi(\lambda)$ is an even, positive, continuous function on $[-\pi, \pi)$, which we think of as the spectral density of the short-memory component of y_t. This impiles that the log-squared returns series becomes a long memory signal plus noise process, $z_t = \log r_t^2 = y_t + w_t$, where the signal y_t is the log-volatility of returns and the noise term $w_t = \log u_t^2 + \log \kappa^2$ is white noise. The persistence in such series, parameterized by the long memory parameter d, has been widely documented empirically, especially using the popular log-periodogram regression (LPR) estimator of Geweke & Porter-Hudak (1983) and Robinson (1995*b*) or the local Whittle (LW) estimator of Künsch (1987) and Robinson (1995*a*). Although the LPR and LW estimators preserve consistency and asymptotic normality when applied to the LMSV model, see Deo & Hurvich (2001) and Arteche (2004), these authors show both theoretically and via simulations that they are heavily biased in that case. Consequently, much effort has been devoted to developing improved methods for the LMSV model, and in particular Sun & Phillips (2003) (for LPR) and Hurvich & Ray (2003) and Hurvich, Moulines & Soulier (2005) (for LW) have proposed estimators that model the additive noise to reduce bias.

In this paper we extend the local Whittle estimator of Hurvich & Ray (2003) by modeling the spectral density of the short-memory component $\phi(\lambda)$ of the log-volatility process as a polynomial (of finite and even order) instead of a constant near frequency zero. Assuming that the added noise process is a martingale difference (as efficient markets theory as well as empirical evidence in section 4.2 below suggests) while the short-memory component of the signal is more dynamic, this approach yields an (order of magnitude) reduction in bias compared to the estimator of Hurvich & Ray (2003) and achieves a faster convergence rate. For pure long memory processes, i.e. processes without the noise w_t, this approach of modeling the (log-)spectral density of the short-memory dynamics by a polynomial was introduced by Andrews & Sun (2004) for the LW estimator, but is novel in the context we examine here.

We show that the estimator is consistent for $d \in (0, 1)$ and asymptotically normal for $d \in (0, 3/4)$ with rate of convergence that is arbitrarily close to the parametric rate, $n^{1/2}$, if the spectral density is sufficiently smooth near frequency zero. We also show that the local polynomial approximation inflates the asymptotic variance of the long memory estimator by a multiplicative constant, but this is clearly off-set (at least in theory) by the faster convergence rate. In a companion paper, Frederiksen, Nielsen & Nielsen (2007), we analyze a local polynomial estimator of the memory parameter in a type of local level or random walk plus noise model, i.e. $z_t = y_t + v_t$, where the noise v_t is allowed to be serially correlated. The proofs of the results in this paper and those in Frederiksen et al. (2007) are very technical but very similar. Therefore, to keep focus here on a description of the properties of the new estimator for the LMSV model along with supporting finite sample results, both simulations and empirical illustrations, we refer to Frederiksen et al. (2007) for proofs.

In a Monte Carlo study we present results to support the new estimator's theoretical properties. The results clearly indicate the importance of the extra flexibility of using a polynomial instead of a constant in approximating the spectral density of the short-memory component near frequency zero. The simulations also show that high bandwidth parameter values are feasible (and may even be preferable) with the new estimator, which is not so for existing methods. Finally, an analysis of daily log-squared exchange rate returns demonstrates the empirical usefulness of the estimator.

The remainder of the paper is organized as follows. Next we introduce the LW approach and formally define the local polynomial Whittle estimator of the LMSV model. In section 3 we establish the properties of the estimator, and section 4 investigates the estimators finite sample performance by simulations and by an empirical study of exchange rate volatility. Section 5 concludes.

2 Local Whittle estimation of d in the LMSV model

Standard local Whittle (LW) estimation of long memory by Künsch (1987) and Robinson (1995a) relies on approximating $\phi(\lambda)$ by a constant, say G, near frequency zero. Based on a sample of size n, the LW estimator is defined as

$$\hat{d}_{LW} = \underset{d \in (-1/2,1]}{\arg\min} \left[\log \hat{G}(d) - 2d \frac{1}{m} \sum_{j=1}^{m} \log \lambda_j \right], \quad \hat{G}(d) = \frac{1}{m} \sum_{j=1}^{m} \lambda_j^{2d} I_y(\lambda_j),$$

where $m = m(n)$ is the bandwidth which tends to infinity as $n \to \infty$ but at a slower rate than n, $\lambda_j = 2\pi j/n$ are the Fourier frequencies, and $I_y(\lambda) = \frac{1}{2\pi n} \left| \sum_{t=1}^{n} y_t e^{it\lambda} \right|^2$ is the periodogram of y_t.

In pure long memory processes, Andrews & Sun (2004) have suggested to model the logarithm of $\phi(\lambda)$ in the vicinity of the origin by a polynomial instead of a constant. Compared to the standard LW estimator, this approach reduces the order of magnitude of the asymptotic bias and achieves a faster rate of convergence, but inflates the asymptotic variance only by a multiplicative constant.

The validity of these LW estimators is to some extent diminished when the observed series is no longer a pure long memory process, but a LMSV model. For the LMSV model, the leading bias term is of order $O(\lambda_m^{2d})$ implying a slower rate of convergence, see Arteche (2004).[2] Thus, we turn next to a modification of the LW estimator that explicitly takes both the additive noise term, w_t, and the short-memory component of y_t into account.

We now impose the following assumption also applied by, e.g., Breidt et al. (1998), Deo & Hurvich (2001), and Arteche (2004), among others.

A1 The processes $\{y_t\}$ and $\{w_t\}$ are independent.

The assumption of independence between the processes $\{y_t\}$ and $\{w_t\}$ rules out the so-called leverage effect, which has been found to be important for stock return volatility (as in, e.g., the FIEGARCH model of Bollerslev & Mikkelsen (1996)) but which does not appear to be relevant for

[2]For the pure long memory process, the leading bias term is $O(\lambda_m^2)$.

the exchange rate returns we examine in our empirical study in section 4. We indicate below how Assumption A1 could be relaxed.

To allow possibly nonstationary volatility as in, e.g., Hurvich & Ray (2003) we generalize (1) to

$$r_t = \begin{cases} \kappa \exp\left(y_t/2\right) u_t & \text{if } d \in (0, 1/2)\,, \\ \kappa \exp\left(\sum_{s=1}^{t} x_s/2\right) u_t & \text{if } d \in [1/2, 1)\,, \end{cases} \tag{3}$$

such that

$$z_t = \begin{cases} y_t + w_t & \text{if } d \in (0, 1/2)\,, \\ \sum_{s=1}^{t} x_s + w_t & \text{if } d \in [1/2, 1)\,, \end{cases} \tag{4}$$

where, if $d \in [1/2, 1)$, x_t has spectrum of the form (2) with memory parameter $d_x = d - 1$. Defining $y_t = \sum_{s=1}^{t} x_s$ if $d \in [1/2, 1)$, this approach allows log-squared returns $z_t = y_t + w_t$ to possibly be nonstationary with memory parameter $d \in (0, 1)$. Since $\{\sum_{s=1}^{t} x_s\}$ is nonstationary, z_t does not have a spectral density if $d \in [1/2, 1)$ but it has a pseudo spectral density, e.g. Hurvich & Ray (1995) and Velasco (1999). Under Assumption A1 the (pseudo) spectral density of z_t in (4) is[3]

$$f_z(\lambda) = \left\{ \begin{array}{ll} f_y(\lambda) + f_w(\lambda) & \text{if } d \in (0, 1/2) \\ \left|1 - e^{i\lambda}\right|^{-2} f_x(\lambda) + f_w(\lambda) & \text{if } d \in [1/2, 1) \end{array} \right\} = \frac{\sigma_w^2}{2\pi}\left(1 + \frac{2\pi\phi(\lambda)}{\sigma_w^2}\lambda^{-2d}\right). \tag{5}$$

In contrast to the standard LW estimator, Hurvich & Ray (2003) explicitly take the added noise term into account by locally (near $\lambda = 0$) fitting the model (5) with $\phi(\lambda)$ approximated by a constant, and they denote this LWN (local Whittle with noise) estimation.

We now apply the bias-reduction idea of Andrews & Sun (2004) and propose to approximate the spectral density of the short-run dynamics, i.e. $\phi(\lambda)$, near frequency zero by a polynomial. Taking λ^{-2d} outside the parenthesis in (5), we thus propose to locally fit the model

$$g(\lambda) = \lambda^{-2d} G\left(1 + h(d, \boldsymbol{\theta}, \lambda)\right) \text{ as } \lambda \to 0, \tag{6}$$

where $h(d, \boldsymbol{\theta}, \lambda) = \Pi(\boldsymbol{\theta}, \lambda) + \theta_{r+1}\lambda^{2d}$ and $\Pi(\boldsymbol{\theta}, \lambda) = \sum_{l=1}^{r} \theta_l \lambda^{2l}$ is an even polynomial of order $2r$ that locally approximates $\phi(\lambda)/\phi(0) - 1$. This allows us to improve on the asymptotic bias (see section 3) incurred by approximating $\phi(\lambda)$ by a constant, i.e. $\Pi(\boldsymbol{\theta}, \lambda) = 0$, as assumed in Hurvich & Ray (2003) and parameterization (P1) of Hurvich et al. (2005).

Concentrating with respect to G, the local polynomial Whittle with noise (LPWN) criterion function for the LMSV model thus becomes

$$L(d, \boldsymbol{\theta}) = \log \hat{G}(d, \boldsymbol{\theta}) + \frac{1}{m}\sum_{j=1}^{m} \log\left(\lambda_j^{-2d}\left(1 + h(d, \boldsymbol{\theta}, \lambda_j)\right)\right), \tag{7}$$

where $\hat{G}(d, \boldsymbol{\theta}) = \frac{1}{m}\sum_{j=1}^{m} \lambda_j^{2d} I_z(\lambda_j) / \left(1 + h(d, \boldsymbol{\theta}, \lambda_j)\right)$. The proposed LPWN estimator is defined as the minimizer of (7) over the admissible set $(d, \boldsymbol{\theta}) \in D \times \Theta$.

[3] To accommodate the leverage effect, Assumption A1 could allow contemporaneous correlation, while the return process remains a martingale difference sequence by replacing y_t with y_{t-1} in (3). An additional assumption of distributional symmetry around $(0, 0)$ would imply that the spectral density decomposition in (5) holds, see Hurvich et al. (2005). Alternatively, the model could be modified along the lines of model (P2) of Hurvich et al. (2005). We do not attempt this here as the leverage effect is presumably not relevant for the exchange rate returns studied below.

3 Properties of the LPWN estimator

Here we introduce the remaining assumptions needed to establish consistency and asymptotic normality of the LPWN estimator for the LMSV model, and consequently we present the main theorem. In the following, true values of the parameters are denoted by subscript zero and $\lfloor x \rfloor$ denotes the integer part of a real number x.

A2 The spectral density of z_t is $f_z(\lambda) = \lambda^{-2d_0} G_0 \frac{\phi(\lambda)}{\phi(0)} + \frac{\sigma_w^2}{2\pi}$, where $\phi(\lambda)$ is a real, even, positive, continuous function on $[-\pi, \pi)$ and $d_0 \in D = [d_1, d_2]$ with $0 < d_1 < d_2 < 1$.

A3 The function $\phi(\lambda)$ is smooth of order s at $\lambda = 0$, where $s > 2r$ and $s \geq 1$. That is, in a neighborhood of $\lambda = 0$, $\phi(\lambda)$ is $\lfloor s \rfloor$ times continuously differentiable with $\lfloor s \rfloor$ −derivative, $\phi^{(\lfloor s \rfloor)}$, satisfying $|\phi^{(\lfloor s \rfloor)}(\lambda) - \phi^{(\lfloor s \rfloor)}(0)| \leq C |\lambda|^{s - \lfloor s \rfloor}$ for some constant $C < \infty$.

Note that Assumption A3 holds for all $s < \infty$ when y_t is a finite order ARFIMA process. The assumptions on $\phi(\lambda)$ are similar to those of Andrews & Sun (2004), and allow us to establish the following Taylor expansion of $\phi(\lambda)$ around $\lambda = 0$ (recall that the odd order derivatives of an even function are zero at frequency zero),

$$\frac{\phi(\lambda)}{\phi(0)} = 1 + \sum_{l=1}^{\lfloor s/2 \rfloor} \theta_l \lambda^{2l} + O(\lambda^s) = 1 + \Pi(\boldsymbol{\theta}, \lambda) + O(\lambda^{\min\{s, 2+2r\}}) \text{ as } \lambda \to 0,$$

where $\theta_l = \frac{1}{(2l)!\phi(0)} \frac{\partial^{2l}}{\partial \lambda^{2l}} \phi(\lambda)|_{\lambda=0}$. Thus, the approximation (6) to (5) is

$$\begin{aligned}
\log\left(f_z(\lambda)/g(\lambda)\right) &= \log\left(\frac{\sigma_w^2}{2\pi\phi(0)}\lambda^{2d} + \frac{\phi(\lambda)}{\phi(0)}\right) - \log\left(1 + h(d, \boldsymbol{\theta}, \lambda)\right) \\
&= \log\left(1 + \frac{O(\lambda^{\min\{s,2+2r\}})}{1 + \Pi(\boldsymbol{\theta}, \lambda) + \theta_{r+1}\lambda^{2d}}\right) \text{ as } \lambda \to 0, \\
\frac{f_z(\lambda)}{g(\lambda)} &= 1 + O(\lambda^{\min\{s,2+2r\}}) \text{ as } \lambda \to 0,
\end{aligned}$$

and the true values of G and $\boldsymbol{\theta}$ are $G_0 = \phi(0)$ and $\boldsymbol{\theta}_0 = (\theta_{0,1}, ..., \theta_{0,r+1})'$, where

$$\theta_{0,l} = \frac{1}{(2l)!\phi(0)} \frac{\partial^{2l}}{\partial \lambda^{2l}} \phi(\lambda)|_{\lambda=0}, l = 1, \ldots, r, \text{ and } \theta_{0,r+1} = \frac{\sigma_w^2}{\phi(0)\, 2\pi}.$$

A4 (a) The signal process y_t has zero mean and admits an infinite order moving average representation $y_t = \sum_{j=0}^{\infty} \alpha_j \varepsilon_{t-j}$ (stationary case) or $\Delta y_t = x_t = \sum_{j=0}^{\infty} \alpha_j \varepsilon_{t-j}$ (nonstationary case), where $\sum_{j=0}^{\infty} \alpha_j^2 < \infty$ and ε_t satisfies, for all t, $E(\varepsilon_t | \mathcal{F}_{t-1}) = 0$, $E(\varepsilon_t^2 | \mathcal{F}_{t-1}) = 1$, $E(\varepsilon_t^3 | \mathcal{F}_{t-1}) = \mu_3 < \infty$, and $E(\varepsilon_t^4 | \mathcal{F}_{t-1}) = \mu_4 < \infty$ almost surely, where $\mathcal{F}_{t-1}$ is the σ-field generated by $\{\varepsilon_s, s < t\}$.

(b) There exists a random variable ε with $E(\varepsilon^2) < \infty$ such that for all $\eta > 0$ and some $K > 0$, $P(|\varepsilon_t| > \eta) < KP(|\varepsilon| > \eta)$.

(c) In a neighborhood of the origin, $\frac{\partial}{\partial\lambda}\alpha(\lambda) = O(|\alpha(\lambda)|/\lambda)$ as $\lambda \to 0$, where $\alpha(\lambda) = \sum_{k=0}^{\infty} \alpha_k e^{ik\lambda}$.

(d) The noise process w_t is zero mean white noise with variance σ_w^2 and such that $\sup_{t\in\mathbb{N}} E(w_t^4) < \infty$. Furthermore, for all $(s,t,r,v) \in \mathbb{N}^4$ such that $s < t$ and $r < v$ we have that $E(w_t w_s w_r w_v) = \sigma_w^4$ if $s = r, t = v$ and zero otherwise.

Since our estimator is a function of the periodogram at nonzero frequencies only, we assume without loss of generality[4] that the signal process y_t has zero mean. Importantly, Assumption A4 allows for non-Gaussian processes. Note that Assumptions A1-A4 imply the assumptions needed on y_t and w_t to prove consistency and asymptotic normality (if, in addition, $d_2 < 3/4$) of the LWN estimator of Hurvich & Ray (2003).

A5 Θ is a compact and convex subset of $\mathbb{R}^{r+1}$ and $\boldsymbol{\theta}_0$ lies in the interior of Θ.

A6 (i) The bandwidth $m = m(n)$ is such that $m^{-1} + mn^{-1} \to 0$.

A6 (ii) The bandwidth $m = m(n)$ is such that $m^{1+4\max\{r,d_0\}} n^{-4\max\{r,d_0\}} \to \infty$ and $m^{\varphi+1/2} n^{-\varphi} \to 0$, where $\varphi = \min\{s, 2+2r\}$.

Note that the two conditions in Assumption A6(ii) are always compatible because $s > 2r$ by Assumption A3. Also note that if $\phi(\lambda)$ is infinitely smooth near frequency zero (e.g. ARFIMA models) then the second condition in Assumption A6(ii) implies that any r can be chosen and the estimator is $n^{1/2-\delta}$ consistent for all $\delta > 0$. In that case the rate of convergence is arbitrarily close to the parametric rate.

The following theorem presents the asymptotic properties of the LPWN estimator. The proof is very similar to the proof in Frederiksen et al. (2007), and is therefore omitted.

Theorem 1 *(i) If Assumptions A1-A5 and A6(i) hold then* $\hat{d} - d_0 = o_P((\log n)^{-5})$.

(ii) If Assumptions A1-A5 and A6(ii) hold and d_0 *lies in the interior of* $D = [d_1, d_2]$ *with* $0 < d_1 < d_2 < 3/4$*, then* $\hat{d}$ *and* $\hat{\boldsymbol{\theta}}$ *are both consistent and*

$$\mathbf{B}_n \begin{pmatrix} \hat{d} - d_0 \\ \hat{\boldsymbol{\theta}} - \boldsymbol{\theta}_0 \end{pmatrix} \xrightarrow{d} N(\mathbf{0}, \boldsymbol{\Omega}_r^{-1}), \quad \boldsymbol{\Omega}_r = \begin{pmatrix} 4 & \boldsymbol{\mu}_r' & \nu \\ \boldsymbol{\mu}_r & \boldsymbol{\Gamma}_r & \boldsymbol{\gamma}_r \\ \nu & \boldsymbol{\gamma}_r' & \omega \end{pmatrix},$$

where $\mathbf{B}_n = \mathbf{B}_n(d_0)$ *is the* $(r+2) \times (r+2)$ *deterministic diagonal matrix with diagonal elements*

$$(\mathbf{B}_n)_{11} = \sqrt{m}, \ (\mathbf{B}_n)_{k+1,k+1} = \sqrt{m}\lambda_m^{2k} \text{ for } k = 1, \ldots, r, \text{ and } (\mathbf{B}_n)_{r+2,r+2} = \sqrt{m}\lambda_m^{2d_0},$$

$\boldsymbol{\mu}_r$ *and* $\boldsymbol{\gamma}_r = \boldsymbol{\gamma}_r(d_0)$ *are r-vectors with k-th element*

$$(\boldsymbol{\mu}_r)_k = -\frac{4k}{(1+2k)^2} \text{ and } (\boldsymbol{\gamma}_r)_k = \frac{4kd_0}{(1+2d_0+2k)(1+2d_0)(1+2k)} \text{ for } k = 1, \ldots, r, \tag{8}$$

[4] In the nonstationary case the zero mean assumption implies that z_t is free of linear trends which entails a loss of generality. However, from an economic viewpoint, deterministic trends in volatility are somewhat artificial.

$\mathbf{\Gamma}_r$ *is the* $r \times r$ *matrix with* (i,k)*-th element*

$$(\mathbf{\Gamma}_r)_{ik} = \frac{4ik}{(1+2i+2k)(1+2i)(1+2k)} \text{ for } i,k = 1,\ldots,r, \tag{9}$$

$\nu = \nu(d_0) = \frac{-4d_0}{(1+2d_0)^2}$, *and* $\omega = \omega(d_0) = \frac{4d_0^2}{(1+4d_0)(1+2d_0)^2}$. *If* $r = 0$ *define* $\mathbf{\Omega}_0 = \begin{pmatrix} 4 & \nu \\ \nu & \omega \end{pmatrix}$.

First of all, we note that by setting $r = 0$ we obtain as a special case the results for the LWN estimator of Hurvich & Ray (2003). Secondly, the leading $(r+1)\times(r+1)$ submatrix of $\mathbf{\Omega}_r$ is the same as that obtained by Andrews & Sun (2004). Third, we note that the asymptotic variance of $\sqrt{m}(\hat{d}-d_0)$ is free of the polynomial parameters $\boldsymbol{\theta}_0$, but it depends on d_0. Moreover, the use of the polynomial $\Pi(\boldsymbol{\theta},\lambda)$ increases the asymptotic variance of $\hat{d}$ by a multiplicative constant compared to the LWN estimator of Hurvich & Ray (2003). Andrews & Sun (2004) obtain a similar result for their polynomial LW estimator in a non-volatility model.

Assumption A6(ii) allows the bandwidth m to be much higher than for the LWN estimator and the standard LW estimator, which require that (assuming $s \geq 2$) $m^5 n^{-4} \to 0$ and $m^{4d_0+1} n^{-4d_0} \to 0$, respectively, see Hurvich & Ray (2003) and Arteche (2004). Thus, Theorem 1 provides an improvement in the rate of convergence relative to existing estimators in the LMSV model. This comes at the cost of an increase in the asymptotic variance by a multiplicative constant, but this is clearly more than off-set by the faster rate of convergence, at least for large n. Moreover, as in Andrews & Sun (2004) we could calculate the asymptotic bias which would be of order $(m/n)^{\varphi}$, where $\varphi = \min\{s, 2+2r\}$, as opposed to $(m/n)^2$ and $(m/n)^{2d_0}$, respectively, for the LWN and LW estimators in Hurvich & Ray (2003) and Arteche (2004). Thus, as in Andrews & Sun (2004) for the pure long memory case, the asymptotic bias has smaller order of magnitude when modeling the spectral density of the short-memory component locally by a polynomial instead of a constant.

4 Finite sample comparison

We now compare the our LPWN estimator with Hurvich & Ray's (2003) LWN estimator in a Monte Carlo simulation study and in an empirical analysis of long memory in exchange rate volatility.

4.1 Monte Carlo simulations

The finite sample performance of the LW estimator in the LMSV model is rather well known, e.g. Hurvich & Ray (2003) and Haldrup & Nielsen (2007), the former of which also demonstrates that the LWN estimator is superior to the LW estimator in terms of bias and RMSE in that case. For this reason, and to conserve space, we only compare the LWN and LPWN (with $r = 1$) estimators.[5]

We generate data according to the LMSV model (3) and (4), i.e. $z_t = y_t + w_t$, where

$$(1-\alpha L)(1-L)^d y_t = (1+\beta L)\eta_t, \quad \eta_t \sim NID(0,\sigma_\eta^2),$$

[5] The results for the LW and polynomial LW estimators are available from the authors upon request.

Table 1: Simulation results with $\alpha = \beta = 0$

		$d = 0.4$				$d = 0.6$			
		LWN		LPWN		LWN		LPWN	
nsr	*n*	Bias	RMSE	Bias	RMSE	Bias	RMSE	Bias	RMSE
Panel A: $m = \lfloor n^{0.7} \rfloor$									
5	1024	0.0210	0.1936	0.0290	0.1808	0.0238	0.1395	0.0355	0.1452
	2048	0.0030	0.1404	-0.0147	0.1381	0.0159	0.1030	0.0242	0.1103
	4096	0.0055	0.1034	-0.0111	0.1004	0.0126	0.0763	0.0171	0.0821
10	1024	0.0103	0.2332	0.0054	0.2181	0.0202	0.1633	0.0048	0.1622
	2048	-0.0003	0.1812	-0.0328	0.1775	0.0164	0.1204	0.0030	0.1161
	4096	0.0067	0.1299	-0.0256	0.1287	0.0133	0.0867	0.0001	0.0868
20	1024	-0.0022	0.2835	-0.0035	0.2658	0.0155	0.1994	-0.0134	0.1979
	2048	-0.0028	0.2365	-0.0436	0.2373	0.0167	0.1431	-0.0192	0.1462
	4096	0.0107	0.1734	-0.0328	0.1750	0.0147	0.1024	-0.0151	0.1073
Panel B: $m = \lfloor n^{0.8} \rfloor$									
5	1024	0.0111	0.1608	-0.0117	0.1532	0.0147	0.1175	-0.0049	0.1272
	2048	0.0016	0.1105	-0.0248	0.1167	0.0168	0.0795	-0.0023	0.0913
	4096	0.0055	0.0791	-0.0130	0.0870	0.0110	0.0576	0.0007	0.0688
10	1024	0.0155	0.2164	-0.0140	0.2056	0.0164	0.1448	-0.0220	0.1510
	2048	0.0027	0.1508	-0.0354	0.1545	0.0185	0.0962	-0.0132	0.1081
	4096	0.0061	0.1014	-0.0228	0.1091	0.0116	0.0682	-0.0076	0.0753
20	1024	0.0076	0.2758	-0.0236	0.2559	0.0161	0.1820	-0.0323	0.1925
	2048	0.0106	0.2238	-0.0300	0.2203	0.0205	0.1199	-0.0184	0.1342
	4096	0.0104	0.1417	-0.0279	0.1492	0.0130	0.0838	-0.0117	0.0920

Notes: Results are based on 10,000 replications. The LPWN estimator is implemented with $r = 1$.

and $w_t = \log u_t^2$ with $u_t \sim NID(0,1)$. Note that the variance of $w_t = \log u_t^2$ is $\sigma_w^2 = \pi^2/2$ regardless of the variance of u_t. Thus, the (long-run) noise-to-signal ratio, $nsr = \frac{\sigma_w^2(1-\alpha)^2}{\sigma_\eta^2(1+\beta)^2}$, is governed by the short-memory parameters α and β and the variance parameter σ_η^2. For each Monte Carlo DGP we generated 1000 artificial time series with 1024, 2048, and 4096 observations[6]. The signal y_t is generated by premultiplying a vector of *i.i.d.* standard normal variates by the Choleski decomposition of the $n \times n$ autocovariance matrix of the desired fractionally integrated process, see also Beran (1994, pp. 215-217). To generate nonstationary series with $d \geq 1/2$, we simulate the ARFIMA process with integration order $d-1$ and cumulate the resulting series. In the maximization of the likelihood functions for the LWN and LPWN estimators, the value of d was constrained to lie in the interval $[0.01, 0.99]$, c.f. Assumption A2, and the polynomial terms θ_{r+1} and $1 + \Pi(\boldsymbol{\theta}, \lambda_j)$ were constrained to be non-negative. If no maximum was found in the interior of the parameter space, we followed Hurvich & Ray (2003) and split the parameter space into several subspaces and started the iterations at the midpoints of each subspace. Then the interior (in its subspace) solution with the best likelihood value was chosen as the estimator. As the starting value for d for the LWN estimator, we used the LW estimate if it was in the interior of $[0.01, 0.99]$ and otherwise we used 0.4, and for the polynomial term we used zero. For the LPWN estimator we used the LWN estimates as starting values for (d, θ_{r+1}) if they were in the interior of the parameter space and otherwise we used $(0.4, 0)$, and for $(\theta_1, ..., \theta_r)$ we used zero. The simulations were run in

[6] The sample sizes are chosen as powers of two in order to use the fast Fourier transform in calculating the periodogram. This speeds up the simulations considerably.

Table 2: Simulation results with $\alpha = 0.8, \beta = 0$

		$d = 0.4$				$d = 0.6$			
		LWN		LPWN		LWN		LPWN	
nsr	n	Bias	RMSE	Bias	RMSE	Bias	RMSE	Bias	RMSE
Panel A: $m = \lfloor n^{0.7} \rfloor$									
5	1024	0.0435	0.1745	0.0444	0.1696	0.0826	0.1519	0.1095	0.1807
	2048	0.0112	0.1142	0.0053	0.1212	0.0562	0.1069	0.0757	0.1398
	4096	-0.0023	0.0847	-0.0039	0.0930	0.0409	0.0763	0.0577	0.1069
10	1024	0.0234	0.2258	0.0012	0.2074	0.0570	0.1635	0.0395	0.1567
	2048	0.0032	0.1624	-0.0317	0.1569	0.0384	0.1171	0.0187	0.1140
	4096	-0.0100	0.1111	-0.0396	0.1159	0.0290	0.0804	0.0145	0.0801
20	1024	-0.0033	0.2695	-0.0113	0.2639	0.0339	0.1943	-0.0110	0.1912
	2048	-0.0018	0.2224	-0.0484	0.2219	0.0247	0.1403	-0.0244	0.1421
	4096	-0.0079	0.1570	-0.0553	0.1668	0.0208	0.0927	-0.0165	0.0981
Panel B: $m = \lfloor n^{0.8} \rfloor$									
5	1024	0.0890	0.1722	0.0210	0.1551	0.1264	0.1664	0.0724	0.1506
	2048	0.0622	0.1134	-0.0019	0.1024	0.1054	0.1288	0.0518	0.1086
	4096	0.0506	0.0834	-0.0057	0.0800	0.0931	0.1071	0.0399	0.0812
10	1024	0.0631	0.2156	0.0010	0.2076	0.0945	0.1640	0.0256	0.1554
	2048	0.0389	0.1368	-0.0245	0.1374	0.0762	0.1190	0.0146	0.1061
	4096	0.0298	0.0919	-0.0228	0.1021	0.0668	0.0912	0.0152	0.0743
20	1024	0.0309	0.2724	-0.0105	0.2646	0.0682	0.1841	-0.0039	0.1909
	2048	0.0296	0.2016	-0.0324	0.2051	0.0539	0.1286	-0.0046	0.1317
	4096	0.0199	0.1299	-0.0324	0.1431	0.0474	0.0894	0.0024	0.0866

Notes: Results are based on 10,000 replications. The LPWN estimator is implemented with $r = 1$.

Ox, see Doornik (2006).

We use the memory parameter values $d = 0.4$ or $d = 0.6$ since it is well documented in the literature (see the references in the introduction) that volatility exhibits long memory and empirically relevant values of the memory parameter are near the stationarity/nonstationarity boundary of 1/2. For the short-run dynamics we choose either $\alpha \in \{-0.8, -0.5, 0, 0.5, 0.8\}, \beta = 0$ or $\alpha = 0, \beta \in \{-0.8, -0.5, 0, 0.5, 0.8\}$. To conserve space we present a subset of the results that correspond most closely to the parameter values found in the empirical study below. For the nsr we choose $nsr \in \{5, 10, 20\}$, and the variance parameter σ_η^2 is set as a function of α and β such that the nsr has the desired value. High values of the nsr are very well documented in the literature, e.g. Breidt et al. (1998) and Hurvich & Ray (2003), and are also supported by our empirical study below. Thus, we in fact put most emphasis on the simulation results with $nsr = 10$ and $nsr = 20$.

Tables 1-4 present the results of the simulations. Generally, higher sample size and higher nsr makes it easier to disentangle the signal and noise components resulting in better estimates. In the presence of short-run dynamics, the LWN estimator generally performs better when the bandwidth is small, whereas the LPWN estimator is also well behaved for the larger bandwidth.

In the case of no short-run dynamics (Table 1) there are no large biases for either estimator, but surprisingly the RMSEs for the two estimators are very similar. When considering autoregressive short-run dynamics in y_t ($\alpha = 0.8, \beta = 0$) in Table 2, the situation for the smaller bandwidth in Panel A is very similar to the case with no short-run dynamics. However, for the higher bandwidth in Panel B the LWN is biased, especially when also $d = 0.6$. On the other hand, the LPWN is

Table 3: Simulation results with $\alpha = 0, \beta = -0.5$

		$d = 0.4$				$d = 0.6$			
		LWN		LPWN		LWN		LPWN	
nsr	n	Bias	RMSE	Bias	RMSE	Bias	RMSE	Bias	RMSE
Panel A: $m = \lfloor n^{0.7} \rfloor$									
5	1024	0.0731	0.2266	0.0363	0.2057	0.0464	0.1499	0.0447	0.1514
	2048	0.0391	0.1622	-0.0026	0.1504	0.0330	0.1094	0.0255	0.1149
	4096	0.0303	0.1123	-0.0006	0.1044	0.0238	0.0778	0.0196	0.0804
10	1024	0.0499	0.2583	0.0200	0.2434	0.0363	0.1671	0.0082	0.1602
	2048	0.0327	0.1991	-0.0189	0.1908	0.0286	0.1230	-0.0022	0.1230
	4096	0.0240	0.1364	-0.0194	0.1356	0.0210	0.0858	0.0024	0.0860
20	1024	0.0081	0.2933	-0.0017	0.2823	0.0224	0.1927	-0.0169	0.1960
	2048	0.0220	0.2459	-0.0266	0.2451	0.0261	0.1471	-0.0200	0.1514
	4096	0.0182	0.1767	-0.0320	0.1827	0.0177	0.0990	-0.0150	0.1051
Panel B: $m = \lfloor n^{0.8} \rfloor$									
5	1024	0.1779	0.2652	0.0364	0.2179	0.0948	0.1537	0.0217	0.1429
	2048	0.1262	0.1817	0.0034	0.1442	0.0724	0.1111	0.0113	0.1011
	4096	0.1009	0.1321	0.0076	0.1016	0.0567	0.0820	0.0089	0.0737
10	1024	0.1290	0.2745	0.0284	0.2604	0.0789	0.1635	0.0063	0.1631
	2048	0.0997	0.1964	-0.0026	0.1848	0.0598	0.1148	0.0003	0.1152
	4096	0.0801	0.1340	0.0006	0.1231	0.0468	0.0828	0.0030	0.0797
20	1024	0.0727	0.3057	0.0127	0.3017	0.0587	0.1820	-0.0089	0.1938
	2048	0.0744	0.2426	-0.0044	0.2369	0.0497	0.1302	-0.0066	0.1378
	4096	0.0651	0.1649	-0.0065	0.1650	0.0376	0.0901	-0.0033	0.0920

Notes: Results are based on 10,000 replications. The LPWN estimator is implemented with $r = 1$.

Table 4: Simulation results with $\alpha = 0, \beta = -0.8$

		$d = 0.4$				$d = 0.6$			
		LWN		LPWN		LWN		LPWN	
nsr	n	Bias	RMSE	Bias	RMSE	Bias	RMSE	Bias	RMSE
Panel A: $m = \lfloor n^{0.7} \rfloor$									
5	1024	0.2302	0.3722	0.1571	0.3200	0.1754	0.2453	0.0976	0.2213
	2048	0.2545	0.3357	0.0693	0.2434	0.1612	0.2031	0.0627	0.1546
	4096	0.2330	0.2678	0.0241	0.1457	0.1258	0.1544	0.0375	0.1054
10	1024	0.1834	0.3521	0.1327	0.3283	0.1579	0.2470	0.0777	0.2310
	2048	0.2106	0.3156	0.0620	0.2612	0.1446	0.1978	0.0443	0.1607
	4096	0.2041	0.2557	0.0273	0.1799	0.1121	0.1478	0.0252	0.1125
20	1024	0.1106	0.3389	0.0870	0.3384	0.1176	0.2466	0.0567	0.2459
	2048	0.1432	0.3141	0.0460	0.2960	0.1192	0.1943	0.0266	0.1811
	4096	0.1667	0.2609	0.0329	0.2304	0.0953	0.1446	0.0137	0.1275
Panel B: $m = \lfloor n^{0.8} \rfloor$									
5	1024	$-^*$	$-^*$	$-^*$	$-^*$	0.2318	0.3135	0.1980	0.2803
	2048	$-^*$	$-^*$	$-^*$	$-^*$	0.2818	0.3053	0.1548	0.2289
	4096	0.1754	0.3969	0.0942	0.2176	0.2647	0.2760	0.0909	0.1560
10	1024	$-^*$	$-^*$	$-^*$	$-^*$	0.2085	0.3041	0.1586	0.2722
	2048	$-^*$	$-^*$	$-^*$	$-^*$	0.2541	0.2854	0.1110	0.2127
	4096	0.1413	0.3786	0.1003	0.2345	0.2379	0.2523	0.0621	0.1489
20	1024	$-^*$	$-^*$	$-^*$	$-^*$	0.1674	0.2861	0.1127	0.2723
	2048	$-^*$	$-^*$	$-^*$	$-^*$	0.2099	0.2565	0.0699	0.2037
	4096	0.0947	0.3544	0.0619	0.2255	0.1993	0.2219	0.0354	0.1343

Notes: Results are based on 10,000 replications. The LPWN estimator is implemented with $r = 1$. An asterisk indicates that, in more than 30% of replications, either there was no convergence or no interior solution was found.

mostly able to completely remove the bias and consequently also has better RMSE in many cases.

In Tables 3 and 4 we present the results for the case with short-run dynamics of the moving average type ($\alpha = 0, \beta = -0.5$ and $\alpha = 0, \beta = -0.8$, respectively). When $\beta = -0.5$, the LWN estimator is biased for the larger bandwidth, whereas the LPWN estimator shows essentially no bias and also has a smaller RMSE than the LWN estimator in many cases. When $\beta = -0.8$, there were some numerical problems optimizing the likelihood functions for a few configurations of the parameters, but for almost all other configurations of nsr, d, n, and m, the LWN estimator is severely biased. The LPWN estimator is in every case in Table 4 able to obtain large reductions in the bias and in many cases able to almost eliminate the bias, especially when the sample size is large. Moreover, the RMSE of the LPWN estimator is clearly superior to that of the LWN estimator for this model, with smaller values in all entries in the table.

Thus, these simulations illustrate the bias-reducing abilities of the new estimator in the LMSV model with empirically relevant values of the long memory parameter, signal-to-noise ratio, and short-run dynamics. Finally, based on the simulations presented here and also the unreported simulations for alternative values of the short-run parameters, α and β, it appears that the higher bandwidth value $m = \lfloor n^{0.8} \rfloor$ may actually be preferable to $m = \lfloor n^{0.7} \rfloor$ for the LPWN estimator in terms of RMSE and often also in terms of bias. The latter finding that a higher bandwidth value may lead to lower bias for the local polynomial estimator is due to the fact that the polynomial parameters, $\boldsymbol{\theta}$, are only consistently estimated if the bandwidth grows sufficiently fast relative to the sample size, see the first term of Assumption A6(ii). So in that sense, high bandwidth values generate better estimates of the polynomial parameters leading to lower bias.[7]

4.2 Long memory in exchange rate volatility

This section analyses empirically the long memory in volatility of daily returns series of DEM/USD, YEN/USD, and USD/GBP exchange rates obtained from the U.S. Federal Reserve Board of Governors H.10 release. The sample covers the period 12/1/1986 – 11/30/2006 for a total of $n = 5,186$ observations.[8] Even though less than 1% of the returns were zero, we based the analysis on adjusted log-squared returns using the method of Fuller (1996, pp. 495-496), i.e. $\log \tilde{r}_t^2 = \log\left(r_t^2 + \kappa\right) - \frac{\kappa}{r_t^2+\kappa}$, where $\kappa = \frac{0.02}{n} \sum_{t=1}^{n} r_t^2$, instead of removing the zero observations.

In Figure 1 we plot the estimated values of d for the three volatility series using the LW, LWN, and LPWN ($r = 1$) estimators. The estimates are shown for a range of relevant values of the bandwidth parameter, $m \in [50, 2000]$, and in the case of the LPWN estimator we include an approximate asymptotic confidence interval given by plus/minus two asymptotic standard errors. Following assumption A6(ii) and the suggestion by Hurvich & Ray (2003), we emphasize the higher

[7] A detailed theoretical investigation of whether this is due to a general theoretical result and of the related issue of optimal bandwidth choice is beyond the scope of this paper. However, the findings in our simulations that higher bandwidths are preferable are in line with some theoretical results for the pure (non-volatility) long memory model, see e.g. Andrews & Sun (2004, sections 6-7) and the references therein.

[8] After the adoption of the Euro on January 1, 1999, the DEM/USD exchange rate has been calculated using the USD/EUR exchange rate and the fixed 1.95583 DEM/EUR exchange rate.

Figure 1: Estimated long memory in exchange rate volatility

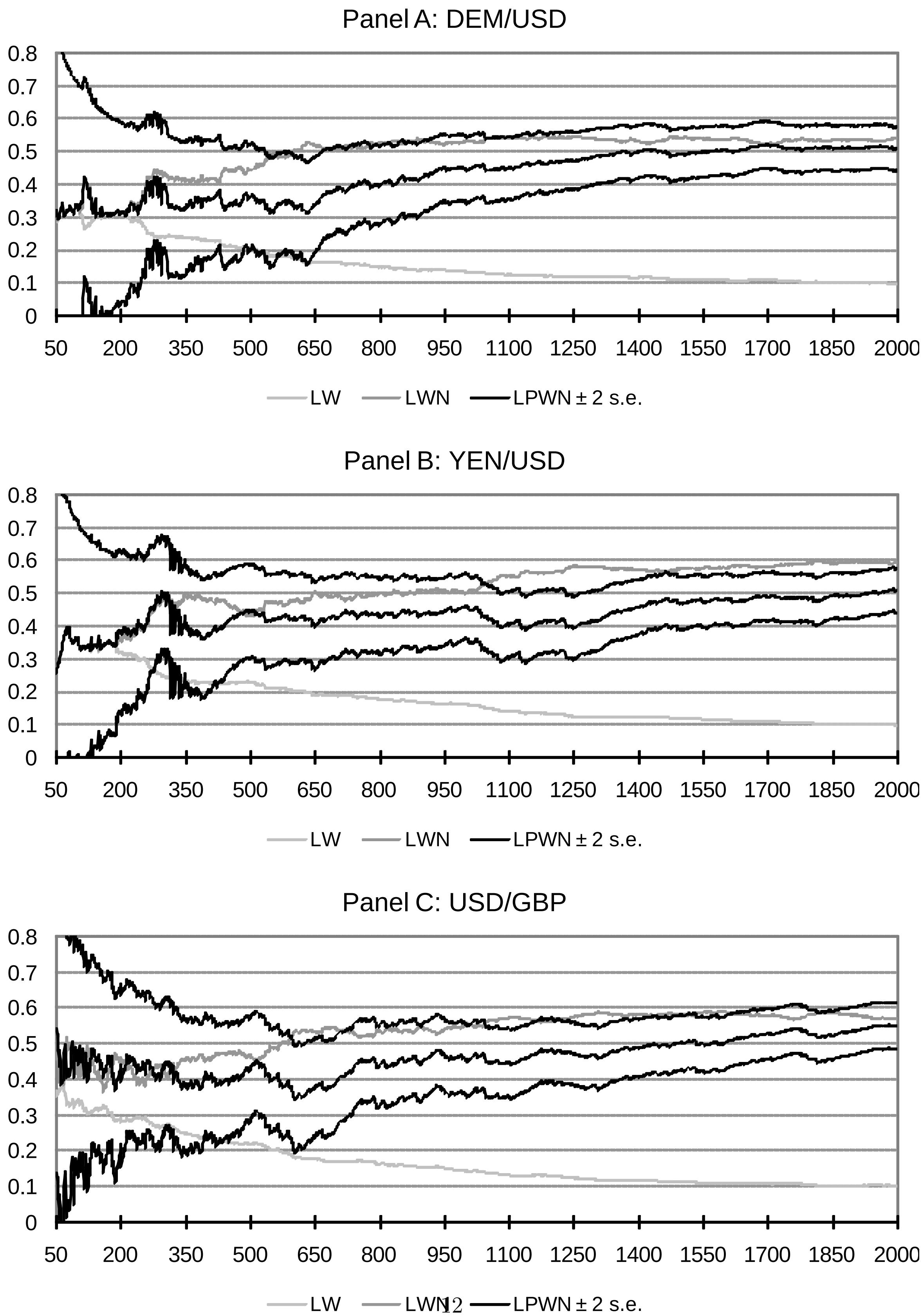

Table 5: Estimated parameters of the parametric volatility model

Exchange rate	$\hat{d}$	$\hat{\alpha}$	$\hat{\beta}$	$\hat{\sigma}_\eta^2$	$\hat{\sigma}_w^2$
DEM/USD	0.4100	0.2875	−0.6466	0.2479	0.2773
	(0.1027)	(0.1081)	(0.0775)	(0.1633)	(0.1597)
YEN/USD	0.4884	−	−0.5087	0.1474	0.3833
	(0.0763)		(0.0921)	(0.0433)	(0.0430)
USD/GBP	0.5687	−	−	0.0138	0.5057
	(0.0749)			(0.0078)	(0.0133)

Note: Standard errors in parentheses.

bandwidth values where the estimates (and confidence intervals) also appear more stable.

The results in Figure 1 show that the LW estimate is smaller than the LWN and LPWN estimates for essentially all bandwidth choices, and the LW estimate is decreasing in the bandwidth. This is expected based on theoretical properties of the LW estimator for the LMSV model. The LWN estimate is higher and shows signs of nonstationarity for higher bandwidth values, whereas the LPWN estimate is in between the LW and LWN estimates. This suggests that the LWN estimate may be upwards biased due to not taking into account the short-run dynamics in the signal process. Indeed, the parametric model below suggests a moving average component with negative coefficient, which according to Table 2 leads to upwards biased LWN estimates when the bandwidth is large and the *nsr* is high. However, the LPWN estimator takes possible short-run dynamics in the signal into account by the polynomial approximation, and hence indicates that the memory parameters for the exchange rate volatility series are not as high as suggested by the LWN estimates.

To stress the importance of the polynomial approximation for the signal process, we also fitted a parametric LMSV-ARFIMA(1,d,1) model to the periodogram of $\log \tilde{r}_t^2$ using the Whittle likelihood, see Fox & Taqqu (1986) and Breidt et al. (1998). The fitted model has spectral density

$$f_z(\lambda) = \frac{\sigma_\eta^2}{2\pi}\,(2\sin\lambda/2)^{-2d}\,\frac{(1+2\beta\cos\lambda+\beta^2)}{(1-2\alpha\cos\lambda+\alpha^2)} + \frac{\sigma_w^2}{2\pi}.$$

We removed insignificant AR and/or MA parameters with resulting estimates reported in Table 5. Note that there is significant short-run dynamics in the signal in two of the three series, and that the estimated (long-run) *nsr*'s, $\frac{\hat{\sigma}_w^2(1-\hat{\alpha})^2}{\hat{\sigma}_\eta^2(1+\hat{\beta})^2}$, are high (4.55, 10.77, and 36.64, respectively).

The high *nsr*'s and significant and negative MA coefficients stress the importance of the LPWN estimator. Thus, looking at the simulation results for high *nsr*'s and negative MA coefficients, the LPWN estimator seems to provide a very useful alternative in empirically relevant scenarios.

5 Concluding remarks

We have proposed a variant of the local polynomial Whittle estimator to semiparametrically estimate the degree of persistence for long memory stochastic volatility models with potential nonstationarity in the volatility process. The estimator is asymptotically normal and can obtain bias reduction as well as a rate of convergence arbitrarily close to the parametric rate, for general short-memory dynamics in the volatility process. A Monte Carlo study supports the theoretical results, and an analysis of daily exchange rates demonstrates the empirical usefulness of the estimators.

References

Andersen, T. G., Bollerslev, T., Diebold, F. X. & Ebens, H. (2001), 'The distribution of realized stock return volatility', *Journal of Financial Economics* **61**, 43–76.

Andersen, T. G., Bollerslev, T., Diebold, F. X. & Labys, P. (2001), 'The distribution of realized exchange rate volatility', *Journal of the American Statistical Association* **96**, 42–55.

Andersen, T. G., Bollerslev, T., Diebold, F. X. & Labys, P. (2003), 'Modelling and forecasting realized volatility', *Econometrica* **71**, 579–625.

Andrews, D. W. K. & Sun, Y. (2004), 'Adaptive local polynomial Whittle estimation of long-range dependence', *Econometrica* **72**, 569–614.

Arteche, J. (2004), 'Gaussian semiparametric estimation in long memory in stochastic volatility and signal plus noise models', *Journal of Econometrics* **119**, 131–154.

Baillie, R. T. (1996), 'Long memory processes and fractional integration in econometrics', *Journal of Econometrics* **73**, 5–59.

Baillie, R. T., Bollerslev, T. & Mikkelsen, H. O. (1996), 'Fractionally integrated generalized autoregressive conditional heteroscedasticity', *Journal of Econometrics* **74**, 3–30.

Beran, J. (1994), *Statistics for Long-Memory Processes*, Chapman-Hall, New York.

Bollerslev, T. & Mikkelsen, H. O. (1996), 'Modeling and pricing long memory in stock market volatility', *Journal of Econometrics* **73**, 151–184.

Breidt, F. J., Crato, N. & de Lima, P. (1998), 'The detection and estimation of long memory in stochastic volatility', *Journal of Econometrics* **83**, 325–348.

Comte, F. & Renault, E. (1998), 'Long memory in continuous-time stochastic volatility models', *Mathematical Finance* **8**, 291–323.

Deo, R. S. & Hurvich, C. M. (2001), 'On the log periodogram regression estimator of the memory parameter in long memory stochastic volatility models', *Econometric Theory* **17**, 686–710.

Ding, Z., Granger, C. W. J. & Engle, R. F. (1993), 'A long memory property of stock returns and a new model', *Journal of Empirical Finance* **1**, 83–106.

Doornik, J. A. (2006), *Ox: An Object-Oriented Matrix Language*, Timberlake Consultants, London.

Fox, R. & Taqqu, M. S. (1986), 'Large-sample properties of parameter estimates for strongly dependent stationary Gaussian series', *Journal of Time Series Analysis* **4**, 221–238.

Frederiksen, P. H., Nielsen, F. S. & Nielsen, M. Ø. (2007), 'Local polynomial Whittle estimation of perturbed fractional processes', *Working paper, Cornell University* .

Fuller, W. A. (1996), *Introduction to statistical time series*, Wiley, New York.

Geweke, J. & Porter-Hudak, S. (1983), 'The estimation and application of long-memory time series models', *Journal of Time Series Analysis* **4**, 221–238.

Haldrup, N. & Nielsen, M. Ø. (2007), 'Estimation of fractional integration in the presence of data noise', *Computational Statistics and Data Analysis* **51**, 3100–3114.

Harvey, A. (1998), Long memory in stochastic volatility, *in* J. Knight & S. Satchell, eds, 'Forecasting Volatility in Financial Markets', Butterworth-Heinemann, London, pp. 307–320.

Henry, M. & Zaffaroni, P. (2003), The long range paradigm for macroeconomics and finance, *in* P. Doukhan, G. Oppenheim & M. S. Taqqu, eds, 'Theory and Applications of Long-Range Dependence', Birkhäuser, Boston, pp. 417–438.

Hurvich, C. M., Moulines, E. & Soulier, P. (2005), 'Estimating long memory in volatility', *Econometrica* **73**, 1283–1328.

Hurvich, C. M. & Ray, B. K. (1995), 'Estimation of the memory parameter for nonstationary or noninvertible fractionally integrated processes', *Journal of Time Series Analysis* **16**, 17–41.

Hurvich, C. M. & Ray, B. K. (2003), 'The local Whittle estimator of long-memory stochastic volatility', *Journal of Financial Econometrics* **1**, 445–470.

Künsch, H. R. (1987), Statistical aspects of self-similar processes, *in* Y. Prokhorov & V. V. Sazanov, eds, 'Proceedings of the First World Congress of the Bernoulli Society', Vol. 1, VNU Science Press, Utrecht, pp. 67–74.

Ray, B. K. & Tsay, R. (2000), 'Long-range dependence in daily stock volatility', *Journal of Business and Economic Statistics* **18**, 254–262.

Robinson, P. M. (1995*a*), 'Gaussian semiparametric estimation of long range dependence', *Annals of Statistics* **23**, 1630–1661.

Robinson, P. M. (1995*b*), 'Log-periodogram regression of time series with long range dependence', *Annals of Statistics* **23**, 1048–1072.

Sun, Y. & Phillips, P. C. B. (2003), 'Nonlinear log-periodogram regression for perturbed fractional processes', *Journal of Econometrics* **115**, 355–389.

Velasco, C. (1999), 'Gaussian semiparametric estimation of non-stationary time series', *Journal of Time Series Analysis* **20**, 87–127.

Research Papers 2008

CREATES
Center for Research in Econometric Analysis of Time Series

2008-22: Mark Podolskij and Daniel Ziggel: A Range-Based Test for the Parametric Form of the Volatility in Diffusion Models

2008-23: Silja Kinnebrock and Mark Podolskij: An Econometric Analysis of Modulated Realised Covariance, Regression and Correlation in Noisy Diffusion Models

2008-24: Matias D. Cattaneo, Richard K. Crump and Michael Jansson: Small Bandwidth Asymptotics for Density-Weighted Average Derivatives

2008-25: Mark Podolskij and Mathias Vetter: Bipower-type estimation in a noisy diffusion setting

2008-26: Martin Møller Andreasen: Ensuring the Validity of the Micro Foundation in DSGE Models

2008-27: Tom Engsted and Thomas Q. Pedersen: Return predictability and intertemporal asset allocation: Evidence from a bias-adjusted VAR model

2008-28: Frank S. Nielsen: Local polynomial Whittle estimation covering non-stationary fractional processes

2008-29: Per Frederiksen, Frank S. Nielsen and Morten Ørregaard Nielsen: Local polynomial Whittle estimation of perturbed fractional processes

2008-30: Mika Meitz and Pentti Saikkonen: Parameter estimation in nonlinear AR-GARCH models

2008-31: Ingmar Nolte and Valeri Voev: Estimating High-Frequency Based (Co-) Variances: A Unified Approach

2008-32: Martin Møller Andreasen: How to Maximize the Likelihood Function for a DSGE Model

2008-33: Martin Møller Andreasen: Non-linear DSGE Models, The Central Difference Kalman Filter, and The Mean Shifted Particle Filter

2008-34: Mark Podolskij and Daniel Ziggel: New tests for jumps: a threshold-based approach

2008-35: Per Frederiksen and Morten Ørregaard Nielsen: Bias-reduced estimation of long memory stochastic volatility