

# Long memory and the relation between implied and realized volatility\*

Federico M. Bandi<sup>†</sup>      Benoit Perron<sup>‡</sup>

*July 2006 (First draft: April 2001)*

## Abstract

We argue that the predictive regression between implied volatility (regressor) and realized volatility over the remaining life of a European option (regressand) is likely to be a fractional cointegrating relation. Since cointegration is associated with long-run comovements, this classical regression cannot be used to test for option market efficiency and short-term unbiasedness of implied volatility as a predictor of realized volatility. Exploiting fractional cointegration and using narrow band spectral methods, we provide consistent estimates of the long-run relation between implied and realized volatility even when implied volatility is measured with error and/or volatility is priced but the volatility risk premium is unobservable. This result is in contrast with much existing work where the use of ordinary least squares methods leads to inconsistent (and downward biased) slope estimates. We find that, while little can be said about short-term unbiasedness, our results largely support a notion of long-run unbiasedness of implied volatility as a predictor of realized volatility.

*JEL classification:* G13, G14, C53.

*Keywords:* Volatility, Predictive Regression, Long Memory.

---

\*This paper was written for the workshop “Modélisation, estimation et prévision de la volatilité/Modeling, estimating and forecasting volatility,” Université de Montréal, April 28, 2001, and was briefly circulated under the title “A fundamentally different interpretation of the relation between implied and realized volatility.” We are grateful to Nagpurnanand Prabhala for providing his option price data. We thank Torben Andersen, Tim Bollerslev, Silvia Gonçalves, Christian Gouriéroux, Javier Hidalgo, Joann Jasiak, Chris Jones, Nour Meddahi, Eric Renault, the conference participants at the 2002 Winter Meetings of the Econometric Society and the 2003 Latin American Meetings of the Econometric Society, and the seminar participants at the University of Southern California for discussions. Bandi acknowledges financial support from the IBM Corporation Faculty Research Fund at the Graduate School of Business of the University of Chicago. Perron acknowledges financial support from the Mathematics of Information Technology and Complex Systems (MITACS) network, the Fonds québécois pour la recherche sur la société et la culture (FQRSC), and the Social Science and Humanities Research Council of Canada (SSHRC).

<sup>†</sup>*Graduate School of Business, The University of Chicago. Address: 5807 South Woodlawn Avenue, Chicago, IL, 60637, USA. Tel. (773) 834-4352. E-mail: federico.band@gsb.uchicago.edu.*

<sup>‡</sup>*Département de sciences économiques, Université de Montréal, CIREQ and CIRANO, C.P. 6128, Succursale centre-ville, Montréal, Québec, H3C 3J7, Canada. Tel. (514) 343-2126. E-mail: benoit.perron@umontreal.ca.*

# 1 Introduction

Implied volatility is widely regarded as the market's best forecast of the future volatility of the asset an option is written on. As a consequence, there has been substantial interest in verifying whether implied volatility is an unbiased forecast of future (ex-post) realized volatility. Consider the regression

$$\sigma_t^R = \alpha + \beta \sigma_t^{BS} + \varepsilon_t, \quad (1)$$

where  $\sigma_t^{BS}$  denotes Black and Scholes implied volatility at time  $t$ ,  $\sigma_t^R$  denotes realized volatility until the option's expiration time  $T$  (with  $T > t$ ), and  $\varepsilon_t$  is a forecast error. Unbiasedness is verified by testing whether the slope and the intercept of this classical regression are equal to 1 and 0, respectively.

The early studies on this topic generally maintained that implied volatility is an informative but biased forecast of future realized volatility. The interested reader is referred to Poteshman (2000) for a complete discussion of the literature. Nonetheless, recent work has pointed out that the evidence in favor of unbiasedness is stronger once the three issues that are believed to contaminate the relevant predictive regression are properly accounted for, namely errors-in-variables (Christensen and Prabhala (1998), Poteshman (2000), and Chernov (2004), among others), overlapping data (Christensen and Prabhala (1998)), and missing variables (Poteshman (2000) and Chernov (2004)). This evidence is generally employed to support option market efficiency and short-term unbiasedness of implied volatility as a predictor of realized volatility over the option's life.

This paper proposes a different interpretation of the classical implied/realized volatility regression. We find that both realized volatility and implied volatility are fractionally integrated (Taylor (1986), Ding et al. (1993), and Dacorogna et al. (1993) are important early references on the relevance of long range dependence in volatility measures, Baillie et al. (1996), Bollerslev and Mikkelsen (1996, 1999), Comte and Renault (1998), Comte et al. (2003), Ray and Tsay (2000), Andersen et al. (2001a), and Andersen et al. (2001b)), among others, provide more recent evidence). More importantly for our purposes, we show that the relation between implied volatility and realized volatility is likely to be a fractionally cointegrating relation.

Since cointegration is associated with long-run comovements, the classical predictive regression of realized volatility on implied volatility should not be regarded as providing information about either option market efficiency or short-term unbiasedness of implied volatility as a predictor of realized volatility. Rather, this regression provides information about the nature of the long-run relation between the two volatility proxies.

Having made these points, we proceed to the estimation of the model. Fractional cointegration provides us with a useful econometric framework to tackle the main issues that pollute standard OLS inference in the implied/realized volatility predictive regression, namely measurement error and risk premia. To this extent, we employ methods that focus on frequencies in the neighborhood of zero and hence on long-run movements, i.e., narrow band spectral least squares (NBLS) methods. Consider the existence of a volatility risk premium (Poteshman (2000) and Chernov (2004)), for example. Traditional semimartingale option pricing implies that the expected value of realized volatility depends (approximately) linearly on implied volatility and a volatility risk premium. Hence, the classical implied/realized volatility predictive regression should include a regressor proxying for the volatility risk premium. Since the risk premium is likely to be correlated with implied volatility, not accounting for it results in inconsistent OLS estimates. The use of NBLS methods allow us to estimate the long-run relation between implied and realized volatility consistently even if a risk premium plays a role. Our consistency result does not depend on a potentially ad-hoc parametrization of the risk premium and the explicit inclusion of the risk premium in the predictive regression, as in Chernov (2004) for example. This is an important feature of our analysis.

Even though short-term unbiasedness cannot be studied, using three data sets (the VIX and the S&P 100 index, S&P 100 option prices and the S&P 100 index, and option prices on Deutschemark-Dollar futures) as well as three commonly-used transformations of the data (variances, standard deviations, and logarithmic standard deviations), we find rather strong empirical support for a long-term notion of the conventional unbiasedness hypothesis.

The paper is organized as follows. Section 2 contains a description of the data and an analysis of the fractional cointegration properties of realized and implied volatility series. In Section 3 we present our econometric approach. Section 4 confirms the validity of our econometric method through simulations relying on an option pricing model with long-memory stochastic

volatility as well as leverage effects in the underlying. The estimated relation between realized and implied volatility is reported in Section 5. Section 6 concludes. The Appendix contains further simulation results and discussions.

## 2 The data

In order to motivate our analysis, we first describe the data and verify the fractional integration and cointegration properties of our volatility measures.

Our main data set consists of observations on the VIX (i.e., the CBOE Market Volatility index) and the S&P 100 (OEX) index from January 1988 to October 2003. Note that the CBOE has recently changed the ticker symbol for the S&P 100 implied volatility to VXO, while VIX now refers to the S&P 500 implied volatility. Throughout the paper, we continue to use the symbol VIX to refer to the S&P 100 implied volatility. The VIX is the implied volatility of a non-traded (synthetic) at-the-money option contract with one month to maturity. This measure is believed to be less affected by the problems that pollute standard implied volatility measures extracted from OEX contracts (i.e., the potential nonsynchronous measurement of option and index levels, early exercise and dividends, bid-ask spreads as well as the wild card option in the OEX market, see Christensen and Prabhala (1998) for discussions). The VIX was first used in predictive regressions of the type analyzed here by Chernov (2004). We use monthly non-overlapping observations to control the correlation structure of the regression errors by taking the closing value of each month, as suggested by Christensen and Prabhala (1998). We also multiply the VIX data by a constant factor equal to  $(\frac{252}{365})^{1/2}$  to account for the difference between trading and calendar days (Schwert (2002)).

As in the early work by French et al. (1987) and Schwert (1989), we compute the realized volatility over the remaining life (one month) of the option as

$$\sigma_t^R = \sqrt{\frac{1}{n_t} \sum_{j=1}^{n_t} r_j^2} \times 252, \quad (2)$$

where  $r_j = \log(S_j/S_{j-1})$ ,  $S_j$  is the closing value of the S&P 100 index on the  $j^{\text{th}}$  trading day of month  $t$  and  $n_t$  is the number of trading days in month  $t$ .

In sum, the data transformations result in two monthly time series,  $\sigma_t^{BS}$  and  $\sigma_t^R$ , where  $\sigma_t^{BS}$  is the annualized (assuming 252 trading days per year) Black and Scholes implied standard

deviation for a synthetic, at-the-money, option contract with one month to maturity, as implied by the VIX, and  $\sigma_t^R$  is the annualized realized standard deviation of the S&P 100 index over the remaining life (one month) of the option, respectively. Both series contain 191 non-overlapping observations.

Table 1 provides descriptive statistics for the two volatility measures.

Table 1 about here

We notice that both implied volatility and realized volatility display heavy tails and positive skewness. As typically the case with volatility measures (see Andersen et al. (2006) for instance), a simple logarithmic transformation (not reported here) would almost lead to normality. The mean of implied volatility is larger than the mean of realized volatility, suggesting a possible risk premium.

We now analyze the persistence properties of the data. As discussed in the Introduction, we model persistence in volatility through a long-memory model. The long-memory parameter  $d$  is estimated in a semiparametric fashion. We report results for both the standard log-periodogram estimator (Geweke and Porter-Hudak (1983), GPH henceforth) and the Gaussian semiparametric local Whittle estimator (Künsch (1987) and Robinson (1995b)). We have also considered the modified log-periodogram estimators of Andrews and Guggenberger (2003) and Sun and Phillips (2003) as well as the modified Whittle estimators of Andrews and Sun (2004) and Hurvich and Ray (2003). Our main findings are not affected by the choice of estimator. The corresponding results are available upon request.

The GPH estimator ( $\hat{d}_{n,m}$ ) is obtained as the OLS estimate of  $d$  in the frequency-domain regression

$$\ln P(\lambda_s) = \gamma_0 - d \ln \left( 4 \left( \sin^2 \left( \frac{\lambda_s}{2} \right) \right) \right) + \varepsilon(\lambda_s), \quad (3)$$

where  $P(\cdot)$  is the periodogram of the data computed at the harmonic frequencies  $\lambda_s = \frac{2\pi s}{n}$  with  $s = 1, \dots, m < \frac{n}{2}$ . The asymptotic standard errors are derived from the result

$$\sqrt{m} \left( \hat{d}_{n,m} - d \right) \Rightarrow \mathbf{N} \left( 0, \frac{\pi^2}{24} \right). \quad (4)$$

Under Gaussian assumptions, the limiting distribution in Eq. (4) was obtained by Robinson (1995a) in the presence of stationary data and by Velasco (1999) in the presence of nonstationary data with  $\frac{1}{2} \leq d < \frac{3}{4}$ . Moreover, for  $\frac{1}{2} < d < 1$  the GPH estimator is known to be consistent (Velasco (1999)). Milder conditions can be invoked for this result to be valid as shown by Kim and Phillips (1999a) and Velasco (2000) using a tapered version of the GPH estimator.

The local Whittle estimator ( $\tilde{d}_{n,m}$ ) of Künsch (1987) and Robinson (1995b) maximizes a frequency-domain Gaussian likelihood for frequencies in the neighborhood of zero, i.e.,

$$\log \left[ \frac{1}{m} \sum_{s=1}^m \lambda_s^{2d} P(\lambda_s) \right] - 2 \frac{d}{m} \sum_{s=1}^m \log \lambda_s.$$

Robinson (1995b) derived the asymptotic behavior of this estimator:

$$\sqrt{m} \left( \tilde{d}_{n,m} - d \right) \Rightarrow \mathbf{N} \left( 0, \frac{1}{4} \right).$$

Since  $1/4 < \pi^2/24 \approx 0.4112$ , the local Whittle estimator is asymptotically more efficient than the GPH estimator for the same choice of  $m$ .

Both of our  $d$  estimators require the use of a bandwidth parameter  $m$  that determines the locality of the estimator. The choice of  $m$  typically involves a standard bias/variance trade-off. Some data-dependent rules have been proposed, for example by Hurvich and Deo (1999) and Andrews and Guggenberger (2003). However, we feel that it is necessary to explore the robustness of our results to the choice of bandwidth. Accordingly, in the figures we report our findings for values of the bandwidth parameter  $m$  between  $[\sqrt{n}]$  and  $[n^{0.8}] + 1$ , where  $[x]$  is the integer part of  $x$ .

In Fig. 1 we report the  $d$  estimates and the corresponding 95% confidence bands for the original volatility series  $\sigma_t^{BS}$  and  $\sigma_t^R$  as well as for their difference (i.e., the residuals obtained by imposing the unbiasedness hypothesis in Eq. (1)) and the OLS residuals from Eq. (1). As we discuss below, there are reasons to believe that the cointegrating vector between implied and realized volatility could be different from  $(1, -1)$ , so the use of the differences should be seen as illustrative. Moreover, since OLS are, in general, consistent only in the nonstationary case ( $d_x \geq 0.5$ ) and since there is uncertainty as to whether the data is stationary or not (see

below), the OLS results should also be seen as illustrative only. For clarity, in Table 2 we report the estimated  $d$  values and standard errors for a few selected bandwidths.

Figure 1 about here

Table 2 about here

In the presence of fractional cointegration, the long-memory parameters of the two volatility series should be equal to each other, whereas the long-memory parameter of the residuals should be less than that of the original volatility series. Fig. 1 provides graphical evidence for this result in our data. As typical in fractional models, the  $d$  parameter estimates are fairly imprecise with large standard error bands (see Baillie (1996)). Nonetheless, these estimates are rather stable over a wide range of bandwidths and similar across volatility measures. Since both sets of residuals tend to have a lower degree of fractional integration, our results suggest cointegration of the fractional type. Interestingly, cointegration in the nonstationary region appears to be possible.

In order to sharpen our estimates of the long memory parameter, Table 2 includes joint estimates constraining the  $d$  parameters for implied volatility and realized volatility to be the same (column 4). We use the multivariate constrained GPH suggested by Robinson (1995a) and the multivariate local Whittle estimator explored by Lobato (1999). Robinson proposes two versions of the multivariate GPH: one that uses OLS to compute the constrained estimator (reported in the table) and one that uses GLS. While the GLS estimator has an asymptotic efficiency advantage over the OLS estimator, this estimator is inconsistent under cointegration since the covariance matrix to be inverted is singular in this case. Similarly, the multivariate Whittle estimator is inconsistent under cointegration (Marinucci and Robinson (2001)). The multivariate GPH OLS estimates tend to be between the corresponding univariate estimates for implied and realized volatility. The local Whittle estimates tend to be lower than either of the univariate estimates.

We have verified the robustness of our finding of fractional cointegration by applying the previous approach to two additional data sets that were employed in previous studies. The first data set was used in Christensen and Prabhala (1998) and consists of monthly data on

the S&P 100 index and OEX options between January 1988 and May 1995 for a total of 89 non-overlapping observations. The data is obtained by examining at-the-money call options with a month to expiration sampled right after the previous expiration date. The second data set consists of monthly data on options on Deutschemark-Dollar futures from the Chicago Mercantile Exchange (CME). The time frame is January 1990 to November 1998 for a total of 107 non-overlapping observations. Similar data is used in Jorion (1995), but the sample there covers the January 1985 to February 1992 period. This market moved from quarterly option expiration dates to monthly expiration dates in 1987. We work with data from 1990 to avoid any potential adjustment/learning effects that this change might have caused. Jorion (1995) discusses the advantages of CME options on foreign currency features over the OEX options. Specifically, empirical work based on the former is expected to be less contaminated by measurement errors since option and underlying trade side-by-side. The results are presented in Figs. 2 and 3 and Tables 3 and 4.

Figure 2 about here

Table 3 about here

Figure 3 about here

Table 4 about here

Our finding of fractional cointegration appears to be a stylized fact. As a final robustness check (not reported here for brevity but available from the authors upon request), we have considered two additional transformations of the data that have been implemented in previous research, namely variances and logarithmic volatilities. Our original formulation, i.e. standard deviations, is used in Canina and Figlewski (1993), Fleming (1998), Jorion (1995), Christensen and Prabhala (1998), and Poteshman (2000), among others. Day and Lewis (1992, 1993), Lamoureux and Lastrapes (1993), and Chernov (2004) use variances. The logarithmic transformation is used in Christensen and Prabhala (1998) and Christensen and Nielsen (2005). The conclusions we draw are consistent with our previous findings.

Having documented fractional cointegration in our data (a fact that necessarily changes our understanding and interpretation of the classical implied/realized volatility predictive regres-



sion), we now turn to an econometric model that accounts for fractional cointegration while providing a solution to the issues that the existing literature has indicated as crucially affecting standard OLS inference in the implied/realized volatility predictive regression.

### 3 Econometric framework

Two main econometric issues have been discussed in the literature on the predictive power of implied volatility, i.e., the presence of an unobserved volatility risk premium and measurement error in implied volatility.

We start with the former. Much recent work has emphasized the importance of explicitly allowing for a time-varying volatility risk premium in option pricing (Benzoni (1998) and Pan (2002), among others) and in the context of the implied/realized volatility regression (Poteshman (2000) and Chernov (2004)). The regression then becomes

$$\sigma_t^R = \alpha^* + \beta^* \sigma_t^{BS} + \gamma RP_t + e_t, \quad (5)$$

where  $RP_t$  denotes a time varying volatility risk premium. In order to explicitly account for a risk premium, the existing work relies on tight parametrizations for it. Unfortunately, simple no-arbitrage principles do not provide us with an expression for the risk premium. If volatility is not a traded asset, markets are incomplete and an infinite number of martingale measures are compatible with the absence of arbitrage. Only a fully-specified equilibrium model could shed some light on the features of the volatility premium (see Pham and Touzi (1996), for example).

From an econometric standpoint, any parametric assumption about the risk premium is, therefore, arbitrary and potentially misleading. In what follows we will account for a risk premium but treat it as unobservable. Interestingly, an unobservable risk premium that is correlated with implied volatility would lead to inconsistent OLS estimates of the parameter  $\beta^*$  in the standard short-memory framework. Exploiting fractional cointegration and using NBLS we are able to estimate the relation between implied and realized volatility *consistently* even when a risk premium correlated with implied volatility plays a role. As we explain below, we do so without having to proxy for it.

Moreover, measurement error in implied volatility could arise from the nonsynchronous measurement of options and index levels as well as the bid-ask spreads in option and underlying prices (see the discussion in Christensen and Prabhala (1998)). Similarly to the presence of an unobserved (and unaccounted for) risk premium correlated with implied volatility, measurement error in implied volatility induces inconsistent OLS slope estimates in the standard short-memory framework. Again, the use of NBLs in our fractionally cointegrated set-up allows us to obtain consistent slope estimates even when implied volatility is measured with noise.

With these observations in mind, we now turn to the estimated model. Denote by  $\mathbf{E}$  the unconditional expectation and define

$$y_t = \sigma_t^R, \quad (6)$$

$$x_t = \sigma_t^{BS} + u_t^{BS}, \quad (7)$$

$$\delta_t = \gamma \{RP_t - \mathbf{E}(RP_t)\}, \quad (8)$$

$$\alpha = \alpha^* + \gamma \mathbf{E}(RP_t), \quad (9)$$

$$\beta = \beta^*, \quad (10)$$

where  $y_t$  and  $x_t$  are the observable volatility series and the shock  $u_t^{BS}$  denotes measurement error in implied volatility. Measurement error in realized volatility, as in Poteshman (2000) and Chernov (2004), can be easily accommodated as well without any effects on our results. Then, Eq. (5) suggests the regression model

$$y_t = \alpha + \beta (x_t - u_t^{BS}) + \delta_t + e_t, \quad (11)$$

or

$$y_t = \alpha + \beta x_t + \varepsilon_t \quad (12)$$

with  $\varepsilon_t = -\beta u_t^{BS} + \delta_t + e_t$ . Consistently with the empirical evidence in Section 2, we model  $x_t$  and  $\varepsilon_t$  as linear processes with spectra satisfying

$$f_x(\lambda) \approx \lambda^{-2d_x} \quad \text{as } \lambda \rightarrow 0^+, \quad (13)$$

$$f_\varepsilon(\lambda) \approx \lambda^{-2d_\varepsilon} \quad \text{as } \lambda \rightarrow 0^+, \quad (14)$$

respectively.

We recognize the fractional cointegration between  $y_t$  and  $x_t$  by assuming that the parameters  $d_x$  and  $d_\varepsilon$  satisfy  $d_x > d_\varepsilon$ . Based on our discussion of the data in Section 2, the fractional parameter  $d_x$  can be either in the stationary region (i.e.,  $d_x < \frac{1}{2}$ ) or in the nonstationary region (i.e.,  $d_x \geq \frac{1}{2}$ ).

Two observations are in order. First, the residuals  $\varepsilon_t$  can be long-range dependent. Even if the measurement error  $u_t^{BS}$  were short-range dependent, the long memory of the residuals may be induced by the quantity  $\delta_t$  (which is a linear function of the de-meaned volatility risk premium). Second, we allow for potential correlation between  $\delta_t$  and the implied volatility proxy  $x_t$  and, as a consequence, between  $\varepsilon_t$  and  $x_t$ . Both properties are important since spot volatility has been documented to be long-range dependent (see Bollerslev and Mikkelsen (1996), among others). If the unobserved risk premium depends on the underlying's spot volatility as often assumed, then the risk premium possesses long memory, and correlation between the implied volatility proxy (the regressor) and the regression residuals containing the risk premium follows immediately.

Notice that contrary to the existing literature, an estimated  $\alpha$  which is significantly different from zero cannot be invoked to rule out “unbiasedness.” In fact, in the presence of a non-vanishing expected risk premium  $\mathbf{E}(RP_t)$  the true  $\alpha$  can be different from zero.

### 3.1 The NBLS estimates

We now turn to the details of the econometric procedure. We notice that Eq. (12) implies

$$y_t - \bar{y} = \beta(x_t - \bar{x}) + (\varepsilon_t - \bar{\varepsilon}), \quad (15)$$

where  $\bar{a} = (\sum_{t=1}^n a_t) / n$  for a generic time-series  $\{a_t : t = 1, \dots, n\}$ . We define the discrete Fourier transforms,

$$\varpi_y(\lambda_s) : = \frac{1}{\sqrt{2\pi n}} \sum_{t=1}^n y_t e^{it\lambda_s} = \frac{1}{\sqrt{2\pi n}} \sum_{t=1}^n (y_t - \bar{y}) e^{it\lambda_s} \quad (16)$$

$$\varpi_x(\lambda_s) : = \frac{1}{\sqrt{2\pi n}} \sum_{t=1}^n x_t e^{it\lambda_s} = \frac{1}{\sqrt{2\pi n}} \sum_{t=1}^n (x_t - \bar{x}) e^{it\lambda_s} \quad (17)$$

$$\varpi_\varepsilon(\lambda_s) : = \frac{1}{\sqrt{2\pi n}} \sum_{t=1}^n \varepsilon_t e^{it\lambda_s} = \frac{1}{\sqrt{2\pi n}} \sum_{t=1}^n (\varepsilon_t - \bar{\varepsilon}) e^{it\lambda_s} \quad (18)$$

at the harmonic (Fourier) frequencies  $\lambda_s = \frac{2\pi s}{n}$ , where  $s = 0, 1, \dots, n-1$ . The (cross-)periodogram, say between  $y$  and  $x$ , is given by

$$P_{yx}(\lambda_s) = \varpi_y(\lambda_s) \overline{\varpi_x(\lambda_s)}. \quad (19)$$

where  $\overline{\varpi_x(\lambda_s)}$  is the complex conjugate of  $\varpi_x(\lambda_s)$ . The NBLs estimator of  $\beta$  is defined as

$$\hat{\beta}_{n,m} = \frac{\operatorname{Re}(\hat{F}_{xy}(1, m))}{\hat{F}_{xx}(1, m)} = \frac{\operatorname{Re}(\hat{F}_{xy}(0, m)) - \bar{x}\bar{y}}{\hat{F}_{xx}(0, m) - \bar{x}^2}, \quad (20)$$

where

$$\hat{F}_{xy}(l, m) = \frac{2\pi}{n} \sum_{s=l}^m P_{yx}(\lambda_s) \quad l = 1, 0 \quad (21)$$

with  $m < \frac{n}{2}$ . In other words, NBLs can be interpreted as OLS in the frequency domain over the first  $m$  frequencies. We will assume that  $m \rightarrow \infty$  as the sample size  $n$  diverges to infinity and  $\frac{m}{n} \rightarrow 0$ , so that we estimate over a degenerating window of frequencies. We estimate  $\alpha$  using

$$\hat{\alpha}_{n,m} = \bar{y} - \hat{\beta}_{n,m} \bar{x}. \quad (22)$$

Some remarks are in order. We start with the stationary case (i.e.,  $d_x < \frac{1}{2}$ ). As discussed by Robinson (1994a), even though OLS estimation is inconsistent in the presence of correlation between the regressor and the residuals, the NBLs estimator of the slope parameter in the same regression is generally consistent provided regressand and regressor are linked by a cointegrating relation. In our framework, correlation between the regressor and the regression error is induced by the likely correlation between  $x_t$  (proxy for implied volatility), the

quantity  $\delta_t$  (which contains the de-measured volatility risk premium) and measurement error. Hence, fractional cointegration in the stationary case allows us to get around the problems posed by the existence of a time-varying volatility risk premium possibly correlated with implied volatility, for example. If the risk premium is of lower order of integration than implied volatility (below we come back to this assumption), then its spectrum is dominated by that of implied volatility near the zero frequency. Performing least squares over a degenerating band of frequencies in the neighborhood of the origin exploits this property. These results are summarized in the following proposition whose proof follows from that of Robinson (1994a) and is therefore omitted.

**Proposition** *Consider the predictive model for realized volatility  $y_t$  given an implied volatility proxy  $x_t$*

$$y_t = \alpha + \beta x_t + \varepsilon_t \tag{23}$$

with  $\varepsilon_t = -\beta u_t^{BS} + \delta_t + e_t$ , where  $u_t^{BS}$  is a measurement error,  $e_t$  is a forecast error and  $\delta_t$  is a linear function of the de-measured time-varying volatility risk premium. Assume  $u_t^{BS}$  and  $\delta_t$  are correlated with  $x_t$ . Also, assume the spectra of  $x_t$  and  $\varepsilon_t$  satisfy

$$f_x(\lambda) \approx \lambda^{-2d_x} \quad \text{as } \lambda \rightarrow 0^+, \tag{24}$$

$$f_\varepsilon(\lambda) \approx \lambda^{-2d_\varepsilon} \quad \text{as } \lambda \rightarrow 0^+, \tag{25}$$

respectively, and  $d_x > d_\varepsilon$  with  $0 \leq d_x, d_\varepsilon < \frac{1}{2}$ . Then, the NBLS estimates defined in Eq. (20) and Eq. (22) are consistent in probability if  $1/m + m/n \rightarrow 0$  as  $n \rightarrow \infty$ .

We now turn to the nonstationary case (i.e.,  $d_x \geq \frac{1}{2}$ ). Typically, in cointegrating relations of the I(1)/I(0) type OLS is consistent even in the presence of regression residuals that are correlated with the regressor since endogeneity only requires second-order adjustments to the asymptotic distribution of the OLS parameter estimates (Park and Phillips (1988)). The same result trivially emerges from fractionally cointegrating regressions of the type analyzed here as shown by Kim and Phillips (1999b) and Robinson and Marinucci (2001, 2003), RM

henceforth, among others. Nonetheless, NBLs methods may entail faster rates of convergence than OLS in some cases (see RM (2001, 2003)), thereby providing justification for their use even in situations where more conventional tools could be employed.

In order to assess the statistical significance of the regression estimates, we rely on subsampling (see Politis et al. (1999) for a complete discussion). We prefer subsampling over the usual bootstrap because of its wider applicability. The only requirements for its validity are the existence of a limiting distribution and conditions limiting the dependence of either the data or the subsampled statistics. For example, subsampling is applicable to the case of an autoregression with a unit root, while the standard bootstrap is not. Another advantage of subsampling over the bootstrap is that the rate of convergence to the asymptotic distribution does not have to be known and can be estimated (see also Bertail et al. (1999)). This property is particularly attractive given that there is uncertainty as to whether we are in the stationary range or not and the convergence rates depend on the long memory parameters of both the regressors and the errors. Moreover, it is likely that the rates of convergence of the constant and slope estimators are different. Our subsampling approach can estimate these different rates consistently. In fact, Politis et al. (1999) use long memory and the different rates of convergence of the sample mean in the stationary and nonstationary case as a motivating example for the need to estimate convergence rates (see their Example 8.1.3).

In our framework, the conditions of Theorems 8.2.2 (consistent estimation of the unknown convergence rate) and 8.3.1 (validity of subsampling with a consistently estimated convergence rate) of Politis et al. (1999) are satisfied under some very standard asymptotic requirements. First, if the parameter of interest is  $\vartheta$ , then  $n^\delta (\hat{\vartheta} - \vartheta^0)$  must have a limiting cumulative distribution function which is continuous and strictly increasing for some  $\delta > 0$ . Virtually all known theoretical results currently available for NBLs take this form, whether the data is stationary or not, as long as the bandwidth is selected to be a fractional power of sample size. Second, the size  $b$  of the subsamples must grow to infinity while the ratio  $b/n$  vanishes as the sample size diverges to infinity. This is accomplished by taking subsamples of size  $b = \lfloor n^{\frac{2}{3}} \rfloor$ , where  $\lfloor x \rfloor$  denotes, as usual, the integer part of  $x$  (see Politis and Romano (1994) and Bertail et al. (1999) for a similar choice), and resampling the couple  $(y_t, x_t)$  in order to preserve the correlation between the regressors and the errors. We compute statistics on all  $n - b + 1$

subsamples of  $b$  successive observations and choose  $m_b$ , the number of Fourier frequencies to compute the NBLs estimator for each subsample, to be the same fraction of the sample size as in the original data, i.e.,  $\frac{m_b}{b} = \frac{m}{n}$ . Under these conditions, we can obtain asymptotically valid confidence intervals without knowing whether we are in the stationary case or in the nonstationary case. In the next section we illustrate the finite sample performance of our subsampling procedure through simulations.

Recently, Christensen and Nielsen (2005), CN henceforth, have provided an important theoretical contribution, namely the asymptotic distribution of the NBLs estimator in the stationary fractionally cointegrated case. Using their theory CN (2005) obtain results similar to ours (in Section 5) about the relation between implied and realized volatility using high-frequency data. Hence, they offer support to the robustness of our results to the use of data sampled at frequencies that are different from the frequencies used in the present study. CN's findings complement the weak convergence results that RM (2001) have derived in the nonstationary fractionally cointegrated situation. For completeness, along with subsampling bands we report in our empirical work both CN's confidence bands (when applicable in the stationary  $d$  case) and RM's confidence bands in the nonstationary  $d$  case.

### 3.2 Interpretation of the slope estimates

Our interest is in estimating the long-run relation between implied and realized volatility over the option's life. We will estimate  $\beta$  and test whether  $\beta = 1$ , thereby implying long-run unbiasedness, but we will not give our  $\beta$  estimates a structural interpretation.

There are two reasons why the true coefficient could be different from one. The first reason has a statistical flavor stemming from our cointegrating framework. The second reason has a superior economic appeal and would be valid even in more conventional settings with short-range dependence (see, for example, Renault (2001)).

We start with the first explanation. In the above Proposition, we have assumed that the risk premium is of lower order of integration than implied volatility. However, it is possible for the risk premium to be of the same order of integration as implied volatility and cointegrated with it. Assume realized volatility is cointegrated with implied volatility and the risk premium (with a cointegrating vector equal to  $(1, -1, -\gamma)$ , say), and the risk premium is cointegrated with

implied volatility (with a cointegrating vector equal to  $(1, -\theta)$ , say), then the true cointegrating vector between realized and implied volatility would not be  $(1, -1)$  but  $(1, -(1 + \gamma\theta))$ . A risk premium that is a linear function of the unobservable spot volatility as in Heston's model (Heston (1993)) would deliver this result. In this case,  $\widehat{\alpha}_{n,m}$  and  $\widehat{\beta}_{n,m}$  would nevertheless be consistent estimates of the parameters of the true long-run relation between implied and realized volatility. However,  $\widehat{\beta}_{n,m}$  would be a consistent estimate of  $(1 + \gamma\theta)$ . Admittedly,  $(1 + \gamma\theta)$  can be smaller or larger than 1 depending on  $\gamma$  and the features of the cointegrating vector between implied volatility and the risk premium as summarized by the coefficient  $\theta$ . If the risk premium depends positively on the level of the underlying's spot volatility and  $\gamma = -1$  (as implied by some option pricing models), then the long-run slope is expected to be smaller than 1.

In sum, our treatment of the risk premium as unobservable leads to two cases: either the risk premium is part of the regression residuals (and does not matter asymptotically) or it is cointegrated with implied volatility and affects the cointegrating vector. Either way, our methods lead to consistent estimates of the true long-run relation between implied and realized volatility series.

We now turn to the second explanation for an estimated slope different from one, namely the existence of a non-zero correlation between innovations to the underlying's returns and innovations to the underlying's spot volatility process, i.e., leverage effects. In the context of a continuous-time model for the underlying with long-memory stochastic spot volatility and leverage  $\rho$ , Comte et al. (2003), CCR henceforth, obtain:

$$\sigma_t^{BS} \approx \left( \sqrt{1 - \rho^2} \right) \mathbf{E}_t (\sigma_t^R) + RP_t, \quad (26)$$

If  $\rho \neq 0$ , as suggested by a substantial amount of recent evidence, CCR's approximation implies that the coefficient of the regression of realized on implied volatility should be larger than one and equal to  $\frac{1}{\sqrt{1 - \rho^2}}$ . Chernov (2004) has recently emphasized that this approximation might be less accurate than the approximation implying  $\beta = 1$ . Hence, the true slope is likely to be somewhere in the interval  $\left[ 1, \frac{1}{\sqrt{1 - \rho^2}} \right]$ . This is consistent with our simulation results reported below.

Of course, both effects could occur at the same time, thereby leading to further complica-



tions which cannot be addressed unless one is willing to impose more structure on the problem. In this paper we do not need to take a stand on the correct  $\beta$  value. Regardless of the true value of  $\beta$ , the interpretation of our regression as an evaluation of the long-run comovement between volatility series remains intact, i.e., long-run unbiasedness can be readily tested.

## 4 Simulations

In this section we evaluate the performance of our econometric approach through simulations. Specifically, we are interested in analyzing the impact of long memory on the predictive regression between implied and realized volatility, the behavior of our proposed NBLs estimator versus OLS, and the performance of our subsampling inference.

We simulate a continuous-time process for the underlying's returns with square-root long-memory stochastic volatility and leverage as in CCR (2003). Using this process, we price a European call option, compute the Black and Scholes implied volatility, and relate implied volatility to the underlying's realized volatility. The simulated model under the risk neutral measure is given by

$$\frac{dS_t}{S_t} = rdt + \sqrt{V_t}dW_t^1, \quad (27)$$

$$V_t = \tilde{v} + In^{(d)}(\tilde{V}_t - \tilde{v}) \quad (28)$$

with

$$d\tilde{V}_t = k(\tilde{v} - \tilde{V}_t)dt + \sigma\sqrt{\tilde{V}_t}(\rho dW_t^1 + \sqrt{1 - \rho^2}dW_t^2), \quad (29)$$

where  $\{W_t^1, W_t^2 : t \geq 0\}$  is a standard Wiener process on the plane and  $In^{(d)}$  is a fractional integration operator used to convert the short-memory spot volatility process  $\tilde{V}_t$  into a long-memory process  $V_t$  with order of fractional integration  $d$  (see CCR (2003) for details). As shown by CCR (2003) both realized volatility and implied volatility display long memory if spot volatility does. Moreover, realized and implied volatility are cointegrated in that case. Hence, CCR's theoretical framework is consistent with our empirical results.

We use the discretized version of the model provided by CCR (2003), with an increment  $\Delta$  of 2 hours (1/12 of a day), to generate sample paths of the underlying price process. We

normalize the logarithmic stock price to be 0 at time 0. We then use the series expansion in Lewis (2000) to compute a sample of 191 implied volatilities given the generated paths of spot volatilities and stock prices. Realized volatility is computed by summing the squares of the daily continuously-compounded returns over the life of the option, i.e., 22 squared returns corresponding to every 12<sup>th</sup> simulated value of the logarithmic price process. We replicate this procedure 1,000 times.

The parameter values are set at those estimated by Ohanissian et al. (2004) using S&P 500 index data, i.e.,  $r = 0.03$ ,  $\tilde{v} = 0.037$ ,  $k = 2.16$ ,  $\sigma = 0.45$ ,  $\rho = -0.48$ , and  $d = 0.34$ . Finally, in order to introduce a likely source of correlation between implied volatility and regression residuals, we run an experiment in which we add measurement error to the implied volatility series. Since our goal is to verify the properties of NBS in the presence of correlation between implied volatility and the regression residuals, adding a measurement error to implied volatility achieves this goal. The measurement error is independent through time and is normally distributed with zero mean and variance that is equal to the variance of the 191 implied volatility realizations over the corresponding replication.

We can easily relate the assumed measurement error in implied volatility to well-known market frictions, such as bid-ask spreads in option prices and bid-ask spreads in the underlying's price, and provide back-of-the-envelope calculations. Given the parameters of the model, the implied volatility mean is 0.1846 in annual terms and its variance is 0.0013. Thus, neglecting nonlinear effects, our measurement error has, on average, a variance of 0.0013. If implied volatility can take only two values with equal probabilities, namely  $0.1846 + c$  and  $0.1846 - c$ , then we need  $c = \sqrt{0.0013} = 0.036$  in order to replicate this type of measurement error. Hence, in this stylized set-up, implied volatility must bounce between 0.1485 and 0.2206. This effect can be accomplished by perturbing any observable variable in the Black and Scholes formula, but we concentrate on the option price and on the price of the underlying. The Black and Scholes price of an at-the-money call option with a month to maturity and a volatility of 18.46% per year with an underlying price of 1 is 0.02306. If implied volatility were 22.06% (14.85%), the option price would be 0.0273 (0.0188). Thus, we need an option bid-ask spread of 0.0085, or about 37% of the Black-Scholes price, to rationalize our measurement error. Figlewski (1997) considers option's bid-ask bounces of up to 25% of the true option price. Alternatively, if we

keep the strike price constant, we need the underlying to bounce between 1.0075 and 0.9916, a total spread of about 1.6% of the efficient price, for implied volatility to have a variance of 0.0013. Figlewski (1997) considers underlying's bid-ask bounces of up to 2% of the efficient price. Our assumed measurement error is purposely large. Our explorations suggests that the relation between size of the measurement error and contaminations in the regression estimates is monotonic. Hence, for brevity, we only report results for the large noise case and the no noise case.

For each replication we estimate the  $d$  parameter for both volatility proxies. In Table 5 we report results of the estimation of the degree of long memory using both the GPH and the Whittle estimator. We now consider 3 bandwidths corresponding to  $n^{0.5}$ ,  $n^{0.6}$ , and  $n^{0.7}$  with  $n = 191$ .

Table 5 about here

In agreement with the existing literature (Andrews and Guggenberger (2003), for example), we find that the  $d$  estimates are consistently upward biased in finite sample. Furthermore, without measurement error both the GPH and the Whittle estimators deliver point estimates for the fractional parameter that are higher for implied than for realized volatility. Finally, as suggested by theory, the Whittle estimator is more precise than the GPH estimator. The addition of measurement error reduces the upward bias of the  $d$  estimates.

For each replication we also run a linear regression of realized volatility on a constant and the implied volatility proxy. This regression is estimated by OLS and NBLs with 6 choices of bandwidth corresponding to powers of the sample size ranging from  $n^{0.25}$  to  $n^{0.8}$ . With our sample size of 191 observations (for the VIX data), these choices amount to including 3, 8, 14, 23, 39, and 66 Fourier frequencies in the relevant regression. We recall that OLS employ all 95 Fourier frequencies.

Table 6 about here

Table 6 contains the means of the quantities of interest (i.e., intercept and slope) across the replications as well as the standard deviations (in parentheses, underneath). The left panel of the table reports results without measurement error, while the right panel provides the same information after adding measurement error to implied volatility. The first column of each panel refers to the estimation of the intercept  $\alpha$ , while the second column refers to the estimation of the slope  $\beta$ . As expected, in the absence of correlation between regressor and disturbance (i.e., without measurement error), OLS provides good results, with a mean estimated slope of about 1.06. NBLs deliver a slightly higher mean slope of about 1.08 when using the three smallest bandwidth choices. The addition of measurement error induces a very strong negative relation between mean estimates and number of frequencies as it introduces correlation between regressor and disturbances. In this case, OLS provides very downward biased estimates in that the slope coefficient is estimated to be only 0.53. NBLs remove most of the bias with an estimated slope above 1.05 for the two smallest bandwidth choices. In other words, the obvious downward trend to the estimates of  $\beta$ , as a function of the number of Fourier frequencies, should be imputed to the accumulation of short-run correlation. To put things in perspective, we recall that, given our adopted parameter values, the slope coefficient implied by CCR's approximate relation is  $\frac{1}{\sqrt{1-\rho^2}} \approx 1.14$ . The slope coefficient given by Chernov's approximate relation is 1. As stressed in Subsection 3.2, we conjecture that the true  $\beta$  value should lay somewhere between 1 and 1.14.

Below the standard deviation of the estimates, we report the frequency that our subsampling confidence intervals (in square brackets) and CN's confidence intervals (in braces) include CCR's theoretical value of 1.14 (we also report the frequency that our subsampling confidence intervals for the intercept contain the value of 0 but no asymptotic intervals are available in this case). The results for the CN's intervals are obtained by using a bandwidth of  $[n^{.7}]$  when computing the values of  $d$  for the regressor and the residuals. There are two numbers in braces: the first one is the percentage of replications where the resulting interval includes 1.14 among the replications where these intervals can be computed. The second number is the percentage of replications where the variance of the estimator could not be computed because the condition  $\hat{d}_x + \hat{d}_\varepsilon < 1/2$  (which is necessary to obtain a positive asymptotic variance in CN's framework) was violated. This phenomenon was quite frequent because of the upward

bias (reported above) in the estimation of  $d$ . We eliminated these replications when computing coverage rates. Overall, the intervals have lower coverage than their nominal level of 95%. The coverage rate of the subsampling intervals is generally higher than the one of CN's intervals. Importantly, the subsampling intervals are less affected by the presence of measurement error, and intervals are also available for the constant (with comparable coverage rates).

To shed light on all of the issues raised by the option pricing simulations, in the Appendix we report results from a different simulation experiment. In order to have complete control of the slope coefficient without invoking approximations, we consider a simple regression framework between two fractionally cointegrated variables with a common order of integration equal to .34. There are three main findings.

First, the presence of strong serial correlation in the underlying noise of the regressor (before integrating to fractional order) induces a large upward bias in the estimation of the regressor's  $d$ . Some experimentation suggests that this strong bias disappears slowly with an increasing sample size. This phenomenon might explain the larger  $d$  estimates obtained here relative to the estimates reported in the realized volatility literature (which uses much larger sample sizes). Similarly, our option-based simulation in this section has strong serial correlation in the underlying noise (the square-root process for  $\tilde{V}_t$ ). This finding explains the upward bias of the reported implied volatility  $d$  estimates in this section.

Second, the regressand has smaller  $d$  estimates than the regressor. Since the regressand (realized volatility in our application) is a linear combination of the regressor (implied volatility in our framework) plus noise, this result is due to measurement error in the regressand. The noise component drives the realized volatility  $d$  estimates downward with respect to the implied volatility  $d$  estimates in the no measurement error case.

Third, and importantly for our purposes, there is little effect of the serial correlation in the regressor's noise on either the regression point estimates or the coverage rates of the confidence intervals. Since in the regression context we control the slope coefficient, we feel comfortable extending this result to the option pricing framework in the present section. As pointed out earlier, we think that the true slope should be between 1 and 1.14 with 1.08 (approximately our NBS estimate in the no measurement error case) being a very likely candidate. This result is interesting and confirms that, with all probability, both CCR's and

Chernov’s approximations could be improved upon in that the true slope lays somewhere between 1 and  $\frac{1}{\sqrt{(1-\rho^2)}}$ . Incidentally, the use of 1.08 as the right slope value would lead to much higher coverage rates than those reported in Table 6.

## 5 The estimated relation between implied and realized volatility

We now turn to the estimation of the model. NBLs estimation of the predictive regression model in Eq. (12) is presented in Fig. 4 as a function of the number of Fourier frequencies included in the spectral regression. We also report the results for a few selected bandwidths in Tables 7-9.

Fig. 4 about here

Table 7 about here

The first row of Fig. 4 corresponds to the results for the S&P 100 index and VIX data (monthly, from January 1988 to October 2003). As in Section 2, we investigate the robustness of our findings by estimating the model in Eq. (12) in the presence of two additional sets of data. The second row corresponds to the results for the S&P 100 index and S&P 100 options (monthly, from January 1988 to May 1995). The last row corresponds to the results for the Deutschemark-Dollar options (monthly, from January 1990 to November 1998). Table 8 and Table 9 are the relevant tables in these cases. Note that the last point on each graph (the point that includes all the Fourier frequencies) refers to the OLS estimate. The left panels contain results for the intercept  $\alpha$ , while the right panels provide results for the slope coefficient  $\beta$ . We also report 95% confidence bands for both sets of parameters based on the subsampling algorithm described in Section 3. Finally, a horizontal line indicates the location of 0 for the intercept and 1 for the slope for ease of interpretation.

We first discuss the S&P 100 - VIX case. As expected, there is a slight downward trend to the slope coefficients (such a trend is more evident when analyzing the other two data sets). This downward trend can be explained by the fact that as the number of frequencies

in the regression increases, a greater proportion of the high-frequency correlation between the error term and the observed implied volatility series is included in the regression, leading to an increased bias of the slope estimates. Nonetheless, the confidence intervals for the slope coefficients include 1 for all choices of frequencies suggesting long-run unbiasedness. As far as the intercept estimates are concerned, their values might point to the existence of a slight (average) long-run compensation for volatility risk.

In addition to the information contained in Fig. 4, Table 7 reports alternative confidence intervals. It includes confidence intervals for a different choice of subsample size ( $[n^{.75}]$ ) as well as confidence intervals using either CN's theory or RM's theory, depending on the estimated  $d$  for each bandwidth: CN's intervals are in braces, while RM's intervals are in angled brackets. Consider, for example, a bandwidth of 3. The NBLs estimate is 1.044 with a subsampling interval between 0.82 and 1.17 for subsamples of size  $[n^{2/3}] = 33$  and between 0.97 and 1.18 for subsamples of size  $[n^{3/4}] = 51$ . The asymptotic intervals are computed by using the joint OLS GPH estimator for  $d_x$  and the GPH estimator for  $d_\varepsilon$ . We report confidence intervals for the same choice of bandwidth used in the estimation of  $d$ . If we chose to estimate  $d_x$  and  $d_\varepsilon$  with  $n^6$  frequencies, the estimate of  $d_x$  would be 0.412. Since this value is in the stationary region, we compute CN's interval. In this case, the residuals have an estimated  $d$  of 0.005, much lower than  $\widehat{d}_x$ . The result is an interval between 0.973 and 1.114, quite close to the subsampling intervals. In most cases, we find the CN's intervals to be pretty close to those based on subsampling, but somewhat narrower. On the other hand, the RM's intervals tend to be quite narrow. Moreover, in this case the distribution of the estimator is entirely on one side of the estimated value. Proposition 6.1 of RM (2001) implies that the direction of the interval is determined by the sign of the covariance between the innovations of the fractionally-differenced residuals and the regressor. In the above case, for example, if we estimate  $d_x$  and  $d_\varepsilon$  with  $\sqrt{n}$  frequencies, we obtain an interval that lies entirely to the left of the point estimate.

Table 8 about here

Table 9 about here

The results pertaining to the other two data sets (bottom two rows of Fig. 4 and Tables 8 and 9) reinforce our previous findings. The subsampling confidence intervals for the slope coefficient include 1 for a fairly broad choice of small number of frequencies. Coherently with a substantial amount of recent work, the use of OLS would lead to rejection of the unbiasedness hypothesis. Furthermore, the downward trend in the point estimates is much more evident. We interpret this finding as evidence that our use of the VIX as the implied volatility measure in our previous data leads to lower measurement error of implied volatility and, consequently, to lower correlation between the regressor and the residuals (i.e., the term  $u_t^{BS}$  is less important), thereby providing more similar slope estimates across different frequencies.

In general, our use of NBLs to remove the effect of the correlation between the residuals and implied volatility is robust to sensible choices of bandwidth. However, it would be desirable to investigate the possibility of a data-based method of bandwidth selection. Robinson (1994b) discusses the choice of optimal bandwidth for spectral estimation with long-memory, but his work would have to be extended to the case of fractional cointegration. Bandwidth selection appears to involve a typical bias/variance trade-off as exemplified by the tendency for the confidence intervals to narrow down as more Fourier frequencies are included in the estimation.

As earlier in Section 2, the results are robust to the choice of the transformation used. Nonetheless, there is a general tendency for the slope estimates based on the logarithmic transformation to be highest, followed by the standard deviation and, lastly, by the variance. The corresponding figures can be provided by the authors upon request.

## 6 Conclusion

We have argued that the persistence of measures of financial market volatility needs to be taken into account when assessing the relation between realized and implied volatility. This paper uses a simple modelling approach based on long memory to capture persistence. In doing so, the features of the data suggest that the usual predictive regression between realized and implied volatility (regressand and regressor, respectively) is, in fact, a fractional cointegrating relation. This finding modifies the interpretation of such regression as a tool to test for option market efficiency, as typically the case in the existing literature, since only long-run comovements between the two volatility series can be investigated.



While the finding of fractional cointegration renders the usual inference carried out in this framework invalid, the potential long memory property of the data allows us to suggest an econometric approach to estimate the relation between realized and implied volatility that is robust to the main issues that were raised in the previous literature as explanations for a slope coefficient less than one, i.e., measurement error and presence of an unobservable time-varying risk premium correlated with the regression residuals. In particular, our approach does not require to choose a particular parametrization for the risk premium and is thus robust to the likely misspecification that this choice would entail. More generally, our procedure is robust to other forms of non-orthogonality between the regressor, the implied volatility proxy, and the regression residuals.

Consistently with some recent studies (Christensen and Prabhala (1998), Poteshman (2000) and Chernov (2004), among others), we do find evidence of unbiasedness of implied volatility as a predictor of realized volatility but interpret the result in terms of presence of a long-run one-to-one comovement between volatility series. Little can be said about short-term unbiasedness, option market efficiency and/or validity of a certain option pricing model, which are the focus of much existing work on the subject.

This paper discusses an instance where the persistence of volatility measures, which is a stylized fact based on our results and much previous work, plays a role in modifying our interpretation and understanding of a well-established economic relation while forcing us to employ appropriate econometric methodologies. More generally, we believe that the persistence of volatility measures should be given more care any time a notion of volatility is employed as a dependent or independent variable in forecasting, as often the case in the empirical finance literature (the study of the risk-return trade-off is a well-known example). Any approach used must be able to capture persistence and account for its implications. This is what our focus on long memory and fractional cointegration accomplishes.

## References

- [1] Andersen, T. G., T. Bollerslev, F. X. Diebold, and P. Labys, "The Distribution of Realized Exchange Rate Volatility," *Journal of American Statistical Association*, 96, 2001a, 42-55.
- [2] Andersen, T. G., T. Bollerslev, F. X. Diebold, and H. Ebens, "The Distribution of Realized Stock Return Volatility," *Journal of Financial Economics*, 61, 2001b, 43-76.
- [3] Andersen, T. G., T. Bollerslev, and F. X. Diebold, "Parametric and Nonparametric Volatility Measurement," in L.P. Hansen and Y. Ait-Sahalia (eds.), *Handbook of Financial Econometrics*, forthcoming. Amsterdam: North-Holland, 2006.
- [4] Andrews, D. W. K. and P. Guggenberger, "A Bias-Reduced Log-Periodogram Regression Estimator for the Long Memory Parameter," *Econometrica*, 71, 2003, 675-712.
- [5] Andrews, D. W. K. and Y. Sun, "Adaptive Local Polynomial Whittle Estimation of Long-Range Dependence," *Econometrica*, 72, 2004, 569-614.
- [6] Baillie, R., "Long Memory Processes and Fractional Integration in Econometrics," *Journal of Econometrics*, 73, 1996, 5-59.
- [7] Baillie, R., T. Bollerslev and H. O. Mikkelsen, "Fractionally Integrated Generalized Autoregressive Conditional Heteroskedasticity," *Journal of Econometrics*, 74, 1996, 3-30.
- [8] Benzoni, L., "Pricing Options under Stochastic Volatility: An Empirical Investigation," Mimeo, 2002.
- [9] Bertail, P., D. N. Politis and J. P. Romano, "On Subsampling with an Unknown Rate of Convergence," *Journal of the American Statistical Association*, 94, 1999, 569-579.
- [10] Bollerslev, T. and H. O. Mikkelsen, "Modeling and Pricing Long-Memory in Stock Market Volatility," *Journal of Econometrics*, 73, 1996, 151-184.
- [11] Bollerslev, T. and H. O. Mikkelsen, "Long-term Equity Anticipation Securities and Stock Market Volatility Dynamics," *Journal of Econometrics*, 92, 1999, 75-99.

- [12] Canina, L. and S. Figlewski, "The Informational Content of Implied Volatility," *The Review of Financial Studies*, 6, 1993, 659-681.
- [13] Chernov, M., "Implied Volatilities as Forecasts of Future Volatility, the Market Risk Premia, and Returns Variability," Mimeo, 2004.
- [14] Christensen, B. J. and M. Nielsen, "Semiparametric Analysis of Stationary Fractional Cointegration and the Implied-Realized Volatility Relation," *Journal of Econometrics*, forthcoming, 2005.
- [15] Christensen, B. J. and N. R. Prabhala, "The Relation Between Implied and Realized Volatility," *Journal of Financial Economics*, 50, 1998, 125-150.
- [16] Comte, F. and E. Renault, "Long-Memory Continuous-Time Models," *Journal of Econometrics*, 73, 1996, 101-149.
- [17] Comte, F. and E. Renault, "Long-Memory in Continuous-Time Stochastic Volatility Models," *Mathematical Finance*, 8, 1998, 291-323.
- [18] Comte, F., L. Coutin, and E. Renault, "Affine Fractional Stochastic Volatility Models," Mimeo, 2003.
- [19] Dacorogna, M. M., U. A. Muller, R. J. Nagler, R. B. Olsen, and O. V. Pictet, "A Geographical Model for the Daily and Weekly Seasonal Volatility in the Foreign Exchange Market," *Journal of International Money and Finance*, 12, 1993, 413-438.
- [20] Day, T. and C. Lewis, "Stock Market Volatility and the Information Content of Stock Index Options," *Journal of Econometrics*, 52, 1992, 267-287.
- [21] Day, T. and C. Lewis, "Forecasting Futures Market Volatility," *Journal of Derivatives*, 1, 1993, 33-50.
- [22] Ding, Z., C. W. J. Granger and R. F. Engle, "A Long Memory Property of Stock Returns and a New Model," *Journal of Empirical Finance*, 1, 1993, 83-106.
- [23] Dittmann, I. and C. W. J. Granger, "Properties of Nonlinear Transformations of Fractionally Integrated Processes," *Journal of Econometrics*, 110, 2002, 113-133.

- [24] Figlewski, S., “Forecasting Volatility,” Mimeo, 1997.
- [25] Fleming, J., “The Quality of Market Volatility Forecasts Implied by S&P 100 Index Option Prices,” *Journal of Empirical Finance*, 5, 1998, 317-354.
- [26] French, K. R., G. W. Schwert, and R. F. Stambaugh, “Expected Stock Returns and Volatility,” *Journal of Financial Economics*, 19, 1987, 3-29.
- [27] Geweke, J. and S. Porter-Hudak, “The Estimation and Application of Long Memory Time Series Models,” *Journal of Time Series Analysis*, 4, 1983, 221-238.
- [28] Heston, S. L., “A Closed-Form Solution for Options with Stochastic Volatility with Applications to Bond and Currency Options,” *Review of Financial Studies*, 6, 1993, 327-344.
- [29] Hurvich, C. M. and B. K. Ray, “The Local Whittle Estimator of Long-Memory Stochastic Volatility,” *Journal of Financial Econometrics*, 1, 2003, 445 - 470.
- [30] Hurvich, C. M. and R. Deo, ”Plug-In Selection of the Number of Frequencies in Regression Estimates of the Memory Parameter of a Long-Memory Time Series,” *Journal of Time Series Analysis*, 20, 1999, 331-341.
- [31] Jorion, P., “Predicting Volatility in the Foreign Exchange Market,” *Journal of Finance*, 50, 1995, 507-528.
- [32] Kim, C. S. and P. C. B. Phillips, “Log Periodogram Regression: The Nonstationary Case,” Mimeo, 1999a.
- [33] Kim, C. S. and P. C. B. Phillips, “Fully Modified Estimation of Fractional Cointegration Models,” Mimeo, 1999b.
- [34] Künsch, H. R., “Statistical Aspects of Self-Similar Processes,” in Prohorov, Y. and V.V. Sazonov (eds), Proceedings of the first World Congress of the Bernoulli Society, Utrecht: VNU Science Press, 1987, 67–74.
- [35] Lamoureux, D. and W. Lastrapes, “Forecasting Stock-Return Variance: Toward an Understanding of Stochastic Implied Volatilities,” *Review of Financial Studies*, 6, 1993, 293-326.

- [36] Lewis, A., *Option Valuation Under Stochastic Volatility*, Finance Press: Newport Beach, CA, 2000.
- [37] Lobato, I., “A Semiparametric Two-Step Estimator for a Multivariate Long Memory Model,” *Journal of Econometrics*, 90, 1999, 129-153.
- [38] Marinucci, D. and P. M. Robinson, “Semiparametric fractional cointegration analysis,” *Journal of Econometrics*, 105, 2001, 225-247.
- [39] Ohanissian, A., J. R. Russell, and R. S. Tsay, “True or Spurious Long Memory in Volatility: Does it Matter for Pricing Options?” Mimeo, 2004.
- [40] Pan, J., “The Jump-Risk Premia Implicit in Options: Evidence from an Integrated Time-Series Analysis,” *Journal of Financial Economics*, 63, 2002, 3-50.
- [41] Park, J. Y. and P. C. B. Phillips, “Statistical Inference in Regressions with Integrated Processes: Part 1,” *Econometric Theory*, 4, 1988, 468-497.
- [42] Pham, H. and N. Touzi, “Intertemporal Equilibrium Risk Premia in a Stochastic Volatility Model,” *Mathematical Finance*, 6, 1996, 215-236.
- [43] Politis, D. N., and J. P. Romano, “Large Sample Confidence Regions Based on Subsamples under Minimal Conditions,” *Annals of Statistics*, 22, 1994, 2031-2050.
- [44] Politis, D. N., J. P. Romano, and M. Wolf, *Subsampling*, Springer-Verlag. New York, 1999.
- [45] Poteshman, A. M., “Forecasting future Volatility from Option Prices,” Mimeo, 2000.
- [46] Ray, B. K. and R. S. Tsay, “Long-Range Dependence in Daily Stock Volatilities,” *Journal of Business and Economic Statistics*, 18(2), 2000, 254-262.
- [47] Renault, E., “Econometric Methods for Option Pricing: Lecture Notes,” Workshop of Financial Mathematics and Econometrics, Montreal, June 26-30, 2001.
- [48] Robinson, P. M., “Semiparametric Analysis of Long-Memory Time Series,” *Annals of Statistics*, 22, 1994a, 515-539.

- [49] Robinson, P. M., “Rates of Convergence and Optimal Spectral Bandwidth for Long Range Dependence,” *Probability Theory and Related Fields*, 99, 1994b, 443-473.
- [50] Robinson, P. M., “Log-Periodogram Regression of Time Series with Long Range Dependence,” *Annals of Statistics*, 23, 1995a, 1048-1072.
- [51] Robinson, P. M., “Gaussian Semiparametric Estimation of Long Range Dependence,” *Annals of Statistics*, 23, 1995b, 1630-1661.
- [52] Robinson, P. M. and D. Marinucci, “Narrow-band Analysis of Nonstationary Processes”, *Annals of Statistics*, 29, 2001, 947-986.
- [53] Robinson, P. M. and D. Marinucci, “Semiparametric Frequency Domain Analysis of Fractional Cointegration”, in Robinson, P. M., ed., *Time Series With Long Memory*, Oxford University Press, 2003.
- [54] Schwert, G. W., “Why Does Stock Market Volatility Change Over Time?” *Journal of Finance*, 44, 1989, 1115-1153.
- [55] Schwert, G. W., “Stock Volatility in the New Millenium: How Wacky is Nasdaq?” *Journal of Monetary Economics*, 49, 2002, 1-24.
- [56] Shimotsu, K. and P. C. B. Phillips, “Exact Local Whittle Estimation of Fractional Integration,” *Annals of Statistics*, 33, 2005, 1890-1933.
- [57] Sun, Y. and P. C. B. Phillips, “Nonlinear Log-periodogram Regression for Perturbed Fractional Processes,” *Journal of Econometrics*, 115, 2003, 355-389.
- [58] Taylor, S., *Modeling Financial Time Series*, Wiley, Chichester, 1986.
- [59] Velasco, C., “Non-Stationary Log-Periodogram Regression,” *Journal of Econometrics*, 91, 1999, 325-371.
- [60] Velasco, C., “Non-Gaussian Log-periodogram Regression,” *Econometric Theory*, 16, 2000, 44-79.

**Table 1. Descriptive statistics**

	Mean	Std. dev.	Skewness	Kurtosis
Implied volatility ( $\sigma^{BS}$ )	.1793	.0608	.89	3.97
Realized volatility ( $\sigma^R$ )	.1606	.0741	1.21	4.35

The table reports descriptive statistics for the realized and implied volatility series for the S&P 100 index and the VIX (the CBOE Market Volatility Index) from January 1988 to October 2003 (191 monthly non-overlapping observations). We report mean, standard deviation, skewness, and kurtosis for both volatility measures. The skewness and kurtosis coefficients are computed after studentizing the relevant quantity, say  $\theta$ , as  $t_\theta = \frac{(\theta_t - \bar{\theta})}{\sigma(\theta)}$ , where  $\sigma(\theta)$  is the standard deviation of  $\theta$ . The two coefficients are equal to  $\frac{1}{n-1} \sum t_\theta^3$  and  $\frac{1}{n-1} \sum t_\theta^4$ , respectively.

**Table 2. Empirical results - Estimates of  $d$  S&P 100 index and the VIX**

	$\hat{d}$ (implied)		$\hat{d}$ (realized)		$\hat{d}$ (common)		$\hat{d}$ (differences)	
	GPH	Whittle	GPH	Whittle	GPH	Whittle	GPH	Whittle
$m = \lceil n^{0.5} \rceil = 13$	.636 (.243)	.642 (.139)	.425 (.243)	.473 (.139)	.530 (.172)	.433 (.098)	.070 (.243)	-.015 (.139)
$\lceil n^{0.6} \rceil = 23$	.437 (.167)	.518 (.104)	.388 (.167)	.445 (.104)	.412 (.118)	.374 (.074)	.074 (.167)	.003 (.104)
$\lceil n^{0.7} \rceil = 39$	.519 (.122)	.521 (.080)	.392 (.122)	.443 (.080)	.455 (.086)	.370 (.057)	.001 (.122)	-.032 (.080)

The table reports long memory parameter estimates for both implied volatility and realized volatility using the Geweke-Porter-Hudak (GPH) estimator and the local Whittle estimator. The number of Fourier frequencies used is equal to the integer part of  $n^{0.5}$ ,  $n^{0.6}$ , and  $n^{0.7}$ . We also report estimates based on joint estimation using the multivariate constrained GPH estimator suggested by Robinson (1995a) and the multivariate local Whittle estimator suggested by Lobato (1999). The last two columns report long memory parameter estimates for the differences  $\sigma^R - \sigma^{BS}$ . Standard deviations are in parenthesis underneath.

**Table 3. Empirical results - Estimates of  $d$   
S&P 100 index and S&P 100 options**

	$\hat{d}$ (implied)		$\hat{d}$ (realized)		$\hat{d}$ (common)		$\hat{d}$ (differences)	
	GPH	Whittle	GPH	Whittle	GPH	Whittle	GPH	Whittle
$m = \lceil n^{0.5} \rceil = 9$	.770 (.318)	.666 (.167)	.545 (.318)	.485 (.167)	.657 (.225)	.349 (.118)	-.437 (.318)	-.393 (.167)
$\lceil n^{0.6} \rceil = 14$	.678 (.233)	.619 (.134)	.634 (.233)	.644 (.134)	.656 (.165)	.461 (.095)	-.045 (.233)	-.185 (.134)
$\lceil n^{0.7} \rceil = 23$	.715 (.171)	.566 (.104)	.573 (.171)	.580 (.104)	.644 (.121)	.475 (.074)	.090 (.171)	.061 (.104)

Note: see Table 2

**Table 4. Empirical results - Estimates of  $d$   
Deutschemark-Dollar futures**

	$\hat{d}$ (implied)		$\hat{d}$ (realized)		$\hat{d}$ (common)		$\hat{d}$ (differences)	
	GPH	Whittle	GPH	Whittle	GPH	Whittle	GPH	Whittle
$m = \lceil n^{0.5} \rceil = 10$	.660 (.293)	.546 (.158)	.598 (.293)	.527 (.158)	.629 (.207)	.517 (.112)	.404 (.293)	.455 (.158)
$\lceil n^{0.6} \rceil = 16$	.530 (.213)	.499 (.125)	.627 (.213)	.589 (.125)	.579 (.150)	.446 (.088)	.273 (.213)	.324 (.125)
$\lceil n^{0.7} \rceil = 26$	.604 (.158)	.561 (.098)	.549 (.158)	.394 (.098)	.576 (.111)	.360 (.069)	.098 (.158)	.031 (.098)

Note: see Table 2.



**Table 5. Simulation results - long memory parameter**

	$\hat{d}$ (implied)				$\hat{d}$ (realized)		$\hat{d}$ (common)			
	without error		with error		GPH	Whittle	without error		with error	
	GPH	Whittle	GPH	Whittle			GPH	Whittle	GPH	Whittle
$m = \lfloor n^{0.5} \rfloor$	.685 (.266)	.671 (.232)	.566 (.235)	.540 (.201)	.609 (.258)	.594 (.225)	.647 (.250)	.375 (.175)	.656 (.268)	.350 (.170)
$m = \lfloor n^{0.6} \rfloor$	.838 (.194)	.842 (.173)	.562 (.145)	.544 (.110)	.642 (.179)	.632 (.146)	.740 (.173)	.494 (.135)	.748 (.177)	.415 (.114)
$m = \lfloor n^{0.7} \rfloor$	.965 (.146)	.976 (.132)	.504 (.096)	.482 (.068)	.598 (.130)	.578 (.103)	.782 (.123)	.612 (.112)	.786 (.123)	.430 (.076)

The table reports long memory parameter estimates for both implied volatility and realized volatility using the Geweke-Porter-Hudak (GPH) estimator and the local Whittle estimator. The number of Fourier frequencies used is equal to the integer part of  $n^{0.5}$ ,  $n^{0.6}$ , and  $n^{0.7}$ . We also report estimates based on joint estimation using the multivariate constrained GPH estimator suggested by Robinson (1995a) and the multivariate local Whittle estimator suggested by Lobato (1999). The table contains the means of the estimates across simulations (with standard deviations in parenthesis underneath). The true parameter is  $d = 0.34$ .

**Table 6. Simulation results - cointegration relation**

	without error			with error		
	$\alpha$	$\beta$		$\alpha$	$\beta$	
OLS	-.011 (.014)	1.059 (.077)		.087 (.017)	.531 (.057)	
$m = \lceil n^{0.25} \rceil = 3$	-.015 (.019) [66.1]	[64.7]	1.082 (.104) {35.1 – 90.3}	-.006 (.027) [77.2]	[74.6]	1.033 (.144) {32.7 – 58.7}
$m = \lceil n^{0.4} \rceil = 8$	-.014 (.015) [67.0]	[66.7]	1.079 (.084) {51.7 – 91.1}	.003 (.020) [89.9]	[89.4]	.988 (.108) {37.7 – 61.5}
NBLS $m = \lceil n^{0.5} \rceil = 13$	-.014 (.015) [67.9]	[66.4]	1.077 (.081) {48.2 – 91.7}	.011 (.018) [91.9]	[93.1]	.944 (.098) {36.9 – 65.6}
$m = \lceil n^{0.6} \rceil = 23$	-.013 (.014) [63.3]	[61.9]	1.073 (.079) {53.2 – 92.1}	.026 (.017) [79.0]	[81.0]	.862 (.085) {24.6 – 77.6}
$m = \lceil n^{0.7} \rceil = 39$	-.013 (.014) [60.0]	[59.9]	1.069 (.078) {55.6 – 91.9}	.045 (.016) [47.9]	[52.0]	.757 (.076) {4.4 – 88.6}
$m = \lceil n^{0.8} \rceil = 66$	-.012 (.014) [59.6]	[60.1]	1.064 (.078) {58.7 – 92.5}	.069 (.016) [12.5]	[13.9]	.627 (.065) {0.0 – 94.3}

The table reports simulation results from running a linear regression between realized volatility and implied volatility. We estimate the model by OLS and NBLS with a number of Fourier frequencies equal to the integer part of  $n^{0.25}$ ,  $n^{0.4}$ ,  $n^{0.5}$ ,  $n^{0.6}$ ,  $n^{0.7}$  and  $n^{0.8}$ . We report the means of the estimates across simulations (with standard deviations in parenthesis underneath). We also report the frequency that a 95% confidence interval computed using subsampling (in square bracket) and Christensen-Nielsen's theory (in braces) includes 1.14. Note that Christensen-Nielsen's theory is only available for the slope parameter. To implement Christensen and Nielsen's asymptotic theory we estimate the long memory  $d_x$  and  $d_\varepsilon$  parameters with GPH and a bandwidth of  $\lceil n^{.7} \rceil$ . The second number in braces is the percentage of replications for which Christensen and Nielsen's variance could not be computed because the condition  $\hat{d}_x + \hat{d}_\varepsilon < 1/2$  (which is necessary to obtain a positive asymptotic variance in Christensen and Nielsen's framework) was violated. Christensen and Nielsen's coverage rates are obtained by excluding those replications.

**Table 7. Empirical results  
S&P 100 index and the VIX**

NBLS bandwidth $m =$	3	10	15	40	95 (OLS)
NBLS point estimate	1.044	1.035	1.037	0.979	0.925
Subsampling 95% CIs					
Block size = 33 ( $[n^{2/3}]$ )	[.823, 1.172]	[.814, 1.160]	[.819, 1.161]	[.858, 1.107]	[.877, 1.197]
51 ( $[n^{3/4}]$ )	[.968, 1.183]	[.956, 1.170]	[.909, 1.195]	[.912, 1.130]	[.931, 1.191]
Asymptotic 95% CIs					
Bandwidth for estimating $d_x$ and $d_\varepsilon$					
$\hat{d}_x = .530$ ( $[n^{0.5}]$ )	$\langle 1.035, 1.042 \rangle$	$\langle 1.017, 1.031 \rangle$	$\langle 1.013, 1.032 \rangle$	$\langle .905, .963 \rangle$	$\langle .766, .890 \rangle$
$\hat{d}_\varepsilon =$	-.016	.002	-.002	.101	.157
$\hat{d}_x = .412$ ( $[n^{0.6}]$ )	{.973, 1.114}	{.965, 1.105}	{.972, 1.103}	-	-
$\hat{d}_\varepsilon =$	.005	.018	.015	.105	.176
$\hat{d}_x = .455$ ( $[n^{0.7}]$ )	{1.007, 1.081}	{.997, 1.074}	{.999, 1.076}	{.917, 1.041}	-
$\hat{d}_\varepsilon =$	-.042	-.033	-.035	.019	.059

The table reports results from running a linear regression between realized volatility and implied volatility. We estimate the model by OLS and NBLS with various bandwidth choices. We report the point estimates in the first row. The other rows report 95% confidence intervals. The intervals in square brackets are based on subsampling with subsamples of size  $[n^{2/3}]$  and  $[n^{3/4}]$ . For each bandwidth choice for estimating  $d_x$  and  $d_\varepsilon$ , we also report intervals based on either Christensen-Nielsen's theory (in braces when  $\hat{d}_x < .5$ ) or Robinson-Marinucci's theory (in angled brackets when  $\hat{d}_x > .5$ ). Some cells are empty (with the symbol - being used) when, despite the fact that  $\hat{d}_x < .5$ , the condition  $\hat{d}_x + \hat{d}_\varepsilon < .5$  is violated. This condition is necessary for the limiting variance of Christensen-Nielsen's distribution to be non-negative.

**Table 8. Empirical results  
S&P 100 index and S&P 100 options**

NBLS bandwidth $m =$	3	10	15	40	44 (OLS)
NBLS point estimate	.888	.909	.870	.702	.696
Subsampling 95% CIs					
Block size = 19 ( $\lceil n^{2/3} \rceil$ )	[.794, 1.230]	[.821, 1.254]	[.774, 1.115]	[.606, .965]	[.640, .941]
28 ( $\lceil n^{3/4} \rceil$ )	[.817, 1.163]	[.860, 1.163]	[.811, 1.049]	[.592, .903]	[.605, .893]
Asymptotic 95% CIs					
Bandwidth for estimating $d_x$ and $d_\varepsilon$					
$\hat{d}_x = .657$ $\hat{d}_\varepsilon =$ $\lceil n^5 \rceil$	$\langle .891, .905 \rangle$ -.281	$\langle .915, .944 \rangle$ -.316	$\langle .877, .915 \rangle$ -.253	$\langle .705, .721 \rangle$ -.125	$\langle .698, .708 \rangle$ -.123
$\hat{d}_x = .656$ $\hat{d}_\varepsilon =$ $\lceil n^6 \rceil$	$\langle .911, 1.024 \rangle$ .074	$\langle .938, 1.091 \rangle$ .048	$\langle .905, 1.087 \rangle$ .095	$\langle .728, .860 \rangle$ .210	$\langle .719, .841 \rangle$ .212
$\hat{d}_x = .644$ $\hat{d}_\varepsilon =$ $\lceil n^7 \rceil$	$\langle .927, 1.137 \rangle$ .172	$\langle .958, 1.215 \rangle$ .154	$\langle .922, 1.201 \rangle$ .186	$\langle .740, .937 \rangle$ .264	$\langle .733, .907 \rangle$ .266

Note: see Table 7.

**Table 9. Empirical results  
Deutschemark-Dollar futures**

NBLS bandwidth $m =$	3	10	15	40	53 (OLS)
NBLS point estimate	1.070	1.118	1.038	.872	.767
Subsampling 95% CIs					
Block size = 22 ( $\lceil n^{2/3} \rceil$ )	[.195, 1.390]	[.552, 1.347]	[.744, 1.192]	[.829, 1.017]	[.715, .927]
33 ( $\lceil n^{3/4} \rceil$ )	[.863, 1.450]	[1.002, 1.375]	[.983, 1.210]	[.851, 1.020]	[.762, .933]
Asymptotic 95% CIs					
Bandwidth for estimating $d_x$ and $d_\varepsilon$					
$\hat{d}_x = .629$ $\lceil n^{.5} \rceil$	$\langle .111, .893 \rangle$	$\langle .403, 1.001 \rangle$	$\langle -.037, .858 \rangle$	$\langle .600, .824 \rangle$	$\langle .771, .796 \rangle$
$\hat{d}_\varepsilon =$	.424	.469	.411	.416	.441
$\hat{d}_x = .579$ $\lceil n^{.6} \rceil$	$\langle 1.086, 1.149 \rangle$	$\langle 1.139, 1.228 \rangle$	$\langle 1.060, 1.145 \rangle$	$\langle .894, .983 \rangle$	$\langle .877, 1.312 \rangle$
$\hat{d}_\varepsilon =$	.226	.199	.247	.361	.425
$\hat{d}_x = .576$ $\lceil n^{.7} \rceil$	$\langle 1.070, 1.072 \rangle$	$\langle 1.122, 1.138 \rangle$	$\langle 1.032, 1.037 \rangle$	$\langle .776, .852 \rangle$	$\langle .563, .723 \rangle$
$\hat{d}_\varepsilon =$	.044	.005	.069	.186	.248

Note: see Table 7.

## Appendix

These simulations are designed to better understand the properties of NBLs. Specifically, we wish to re-create an environment where (1) the slope of the fractional cointegrating regression is known and (2) the underlying shocks  $u_t$  have different persistence properties. Both features are important but cannot be considered in the context of the option simulations in the main text. In that case, in fact, the true slope of the fractional cointegrating regression can only be approximated (as in CCR (2003) and Chernov (2004), for example) but is unknown in practise. In addition, simulating using shocks (the  $V_t$ 's in the main text) with different properties would take us outside of the realm of realistic processes for the underlying, i.e.,  $V_t$  is known to be highly persistent. We simulate the following DGP:

$$\begin{aligned} y_t &= x_t^* + \varepsilon_t, \\ x_t^* &= (1 - L)^{-d} u_t; u_t = \rho u_{t-1} + v, \\ \begin{pmatrix} \varepsilon_t \\ v_t \end{pmatrix} &\sim iidN(0, I_2), \\ d &= 0.34, \end{aligned}$$

Case 1:  $x_t = x_t^*$  (no measurement error)

Case 2:  $x_t = x_t^* + \sigma(x_t^*) \eta_t$ ;  $\eta_t \sim iidN(0, 1)$  (measurement error).

Our main findings from the simulations in this Appendix are:

- [1] Without serial correlation in the  $u_t$ 's and measurement error, both the GPH and the local Whittle estimator of  $d$  are nearly unbiased. The local Whittle estimator is always more efficient. The presence of measurement error lowers the estimates of  $d$  relative to the same case without error. Similarly, estimation of  $d$  using the regressand produces  $d$  estimates that are lower than those obtained by using the regressor without error. This is because the regression errors  $\varepsilon_t$  play the same role as classical measurement error.
- [2] Serial correlation in the  $u_t$ 's biases the estimates of  $d$  upwards dramatically. Whittle is a little less biased and more efficient than GPH.
- [3] Contrary to the univariate estimators, joint estimation is almost not affected by measurement error. Joint GPH estimation does not lead to substantial efficiency gains. Joint Whittle leads to downward biased estimates.
- [4] Without measurement error, all slope and intercept estimators are unbiased. Precision improves with the number of Fourier frequencies in the regression (OLS tends to be the estimator with the smallest variance).
- [5] The presence of measurement error induces a negative relation between the number of Fourier frequencies and the point estimates (OLS is the most biased estimator). NBLs removes some, but not all, of the downward bias. Estimation is less precise than without measurement error.
- [6] With serial correlation in the underlying  $u_t$ 's, the slope estimators tend to be more precise without much effect on their bias while the intercept estimators are biased. The presence of both measurement error and serial correlation makes the intercept estimators highly biased and less precise as we increase the number of Fourier frequencies in the regression.
- [7] Overall, the subsampling intervals provide the most robust inference across the four cases. The presence of serial correlation has little effect on them, but the presence of measurement error deteriorates their coverage rate, especially for larger bandwidths since they do not fully capture the bias in the estimator.
- [8] Christensen and Nielsen's asymptotic intervals perform best in the simplest case (no serial correlation, no measurement error). With serial correlation, the intervals can almost never be computed because of the upward bias in the estimation of  $d$ .

Table A1. Simulation results for the regression model - long-memory parameter

	$\hat{d}_x$				$\hat{d}_y$		$\hat{d}$ (common)			
	without error		with error		GPH	Whittle	without error		with error	
	GPH	Whittle	GPH	Whittle			GPH	Whittle	GPH	Whittle
$\rho = 0$										
$m = \lceil n^{0.5} \rceil$	.355 (.240)	.324 (.209)	.265 (.242)	.237 (.199)	.280 (.234)	.251 (.202)	.312 (.216)	.181 (.152)	.319 (.226)	.158 (.147)
$m = \lceil n^{0.6} \rceil$	.349 (.166)	.328 (.141)	.242 (.162)	.224 (.133)	.263 (.166)	.241 (.136)	.306 (.152)	.189 (.105)	.307 (.156)	.157 (.100)
$m = \lceil n^{0.7} \rceil$	.348 (.118)	.328 (.096)	.220 (.116)	.206 (.092)	.241 (.120)	.223 (.095)	.295 (.107)	.195 (.075)	.294 (.108)	.150 (.071)
$\rho = 0.95$										
$m = \lceil n^{0.5} \rceil$	1.057 (.244)	1.065 (.208)	.772 (.201)	.756 (.158)	1.046 (.249)	1.052 (.206)	1.052 (.244)	.662 (.191)	1.079 (.258)	.660 (.193)
$m = \lceil n^{0.6} \rceil$	1.132 (.178)	1.138 (.146)	.638 (.124)	.625 (.089)	1.091 (.177)	1.094 (.144)	1.112 (.173)	.766 (.139)	1.130 (.176)	.744 (.135)
$m = \lceil n^{0.7} \rceil$	1.171 (.141)	1.164 (.113)	.500 (.095)	.496 (.057)	1.068 (.128)	1.049 (.101)	1.120 (.128)	.832 (.107)	1.130 (.128)	.766 (.100)

The table reports long memory parameter estimates for both regressor and regressand using the Geweke-Porter-Hudak (GPH) estimator and the local Whittle estimator. The number of Fourier frequencies used is equal to the integer part of  $n^{0.5}$ ,  $n^{0.6}$ , and  $n^{0.7}$ . We also report estimates based on joint estimation using the multivariate constrained GPH estimator suggested by Robinson (1995a) and the multivariate local Whittle estimator suggested by Lobato (1999). The table contains the means of the estimates across simulations (with standard deviations in parenthesis underneath). The true parameter is  $d = 0.34$ .

Table A2. Simulation results for the regression model - regression estimates

		without measurement error ( $\eta_t = 0$ )		with measurement error ( $\eta_t \sim N(0, 1)$ )	
		$\alpha$	$\beta$	$\alpha$	$\beta$
OLS		-.000 (.075)	.998 (.065)	-.007 (.204)	.501 (.055)
$\rho = 0$	$m = \lceil n^{0.25} \rceil = 3$	.002 (.099) [81.5]	1.005 (.192) [77.3] {77.7 - 9.5}	-.005 (.144) [79.1]	.857 (.245) [74.6] {75.6 - 4.7}
	$m = \lceil n^{0.4} \rceil = 8$	.001 (.085) [85.9]	.996 (.130) [86.5] {84.1 - 9.3}	-.004 (.135) [83.3]	.788 (.152) [77.4] {62.0 - 5.4}
	NBLS $m = \lceil n^{0.5} \rceil = 13$	.001 (.082) [86.4]	.995 (.112) [90.0] {86.1 - 9.4}	-.005 (.141) [81.2]	.745 (.126) [68.0] {43.8 - 5.5}
	$m = \lceil n^{0.7} \rceil = 39$	.000 (.078) [81.4]	.995 (.084) [81.8] {88.2 - 10.8}	-.006 (.171) [69.3]	.617 (.081) [3.3] {5.0 - 12.7}
OLS		-.002 (.099)	1.000 (.011)	-.064 (3.395)	.502 (.028)
$\rho = .95$	$m = \lceil n^{0.25} \rceil = 3$	-.001 (.107) [88.2]	1.001 (.014) [87.0] {100.0 - 99.9}	-.004 (.778) [85.6]	.976 (.081) [84.3] {87.0 - 94.6}
	$m = \lceil n^{0.4} \rceil = 8$	-.002 (.101) [90.6]	1.000 (.012) [91.2] {100.0 - 99.9}	.001 (.836) [93.1]	.930 (.065) [94.8] {88.9 - 95.6}
	NBLS $m = \lceil n^{0.5} \rceil = 13$	-.002 (.099) [91.1]	1.000 (.012) [91.6] {100.0 - 99.9}	-.005 (1.002) [91.8]	.887 (.058) [91.6] {78.3 - 97.7}
	$m = \lceil n^{0.7} \rceil = 39$	-.002 (.099) [89.1]	1.000 (.012) [87.8] {100.0 - 99.9}	-.037 (2.010) [85.0]	.713 (.040) [64.3] {0.0 - 99.9}



The table reports simulation results from running a linear regression between  $y_t$  and  $x_t$ . We estimate the model by OLS and NBLs with a number of Fourier frequencies between  $[n^{0.25}]$  and  $[n^{0.7}]$ . We report the means of the estimates (with standard deviations in parenthesis underneath). We also report the frequency that a 95% confidence interval computed using subsampling (in square bracket) and Christensen-Nielsen's theory (in braces), includes the true slope of 1. Note that Christensen-Nielsen's theory is only available for the slope parameter. To implement Christensen and Nielsen's asymptotic theory we estimate the long memory  $d_x$  and  $d_\varepsilon$  parameters with GPH and a bandwidth of  $[n^7]$ . The second number in braces is the percentage of replications for which Christensen and Nielsen's variance could not be computed because the condition  $\hat{d}_x + \hat{d}_\varepsilon < 1/2$  (which is necessary to obtain a positive asymptotic variance in Christensen and Nielsen's framework) was violated. Christensen and Nielsen's coverage rates are obtained by excluding those replications.

Figure 1. d estimates for S&P 100 and VIX data

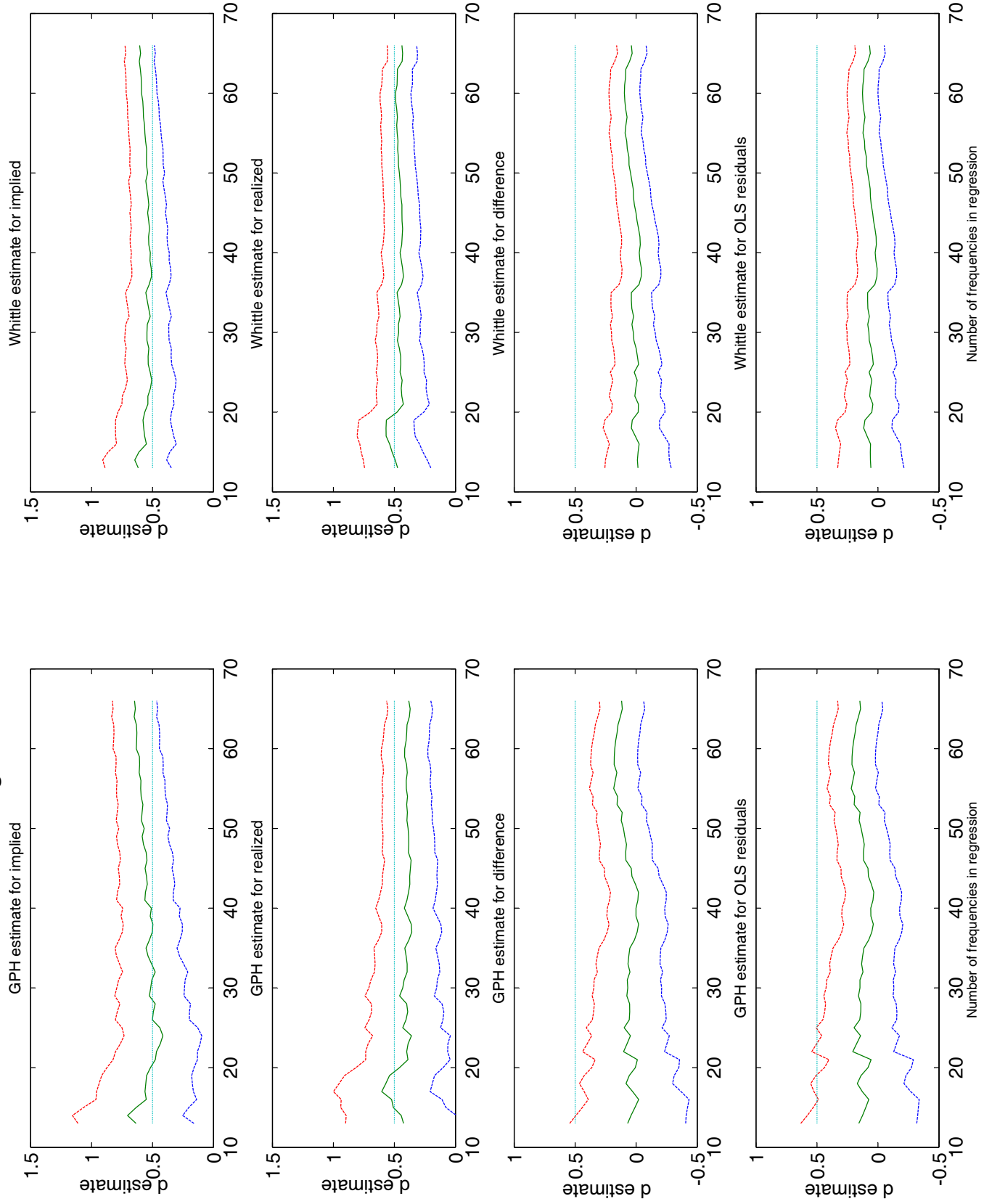


Figure 2. d estimates for S&P 100 data

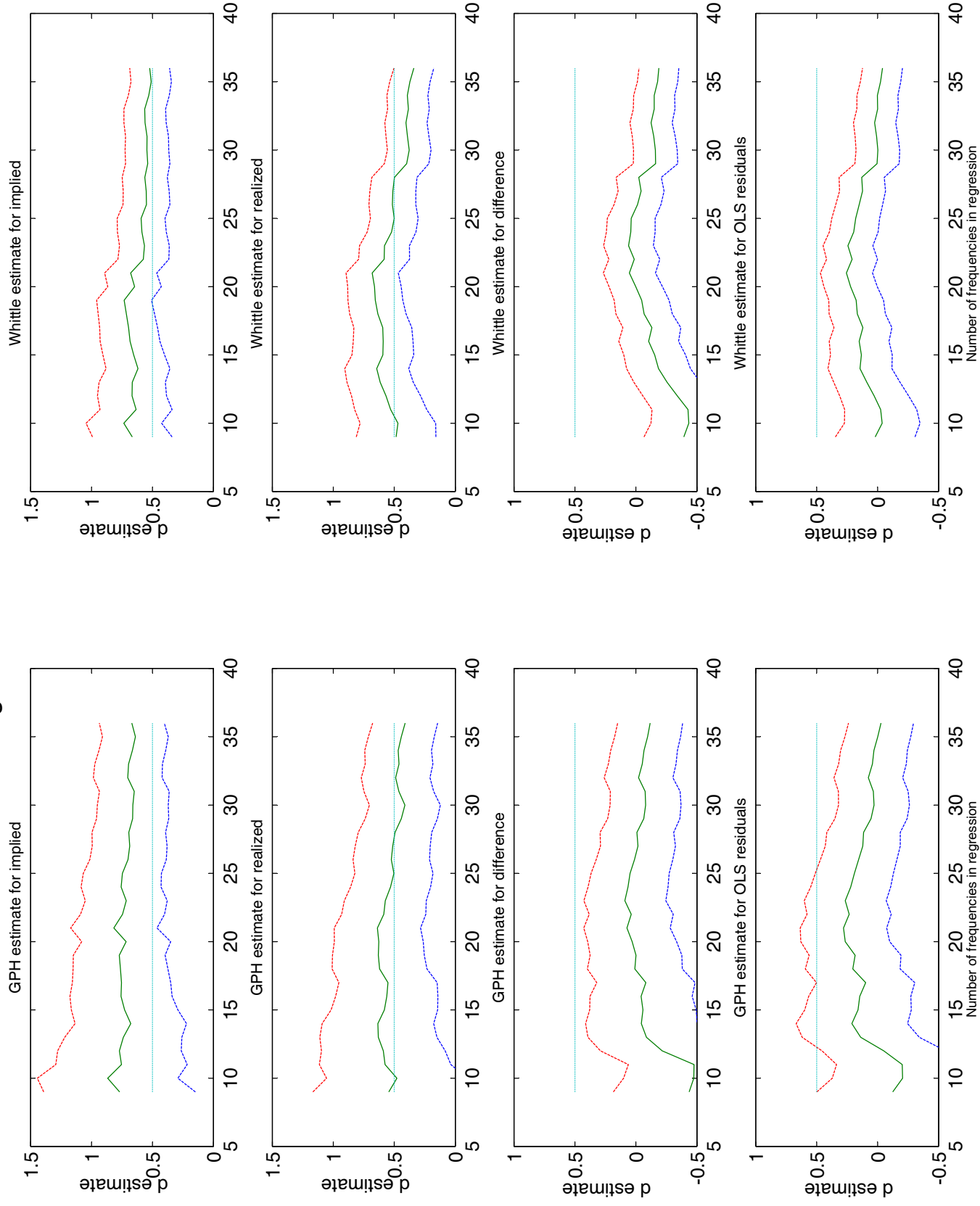


Figure 3. d estimates for DM data

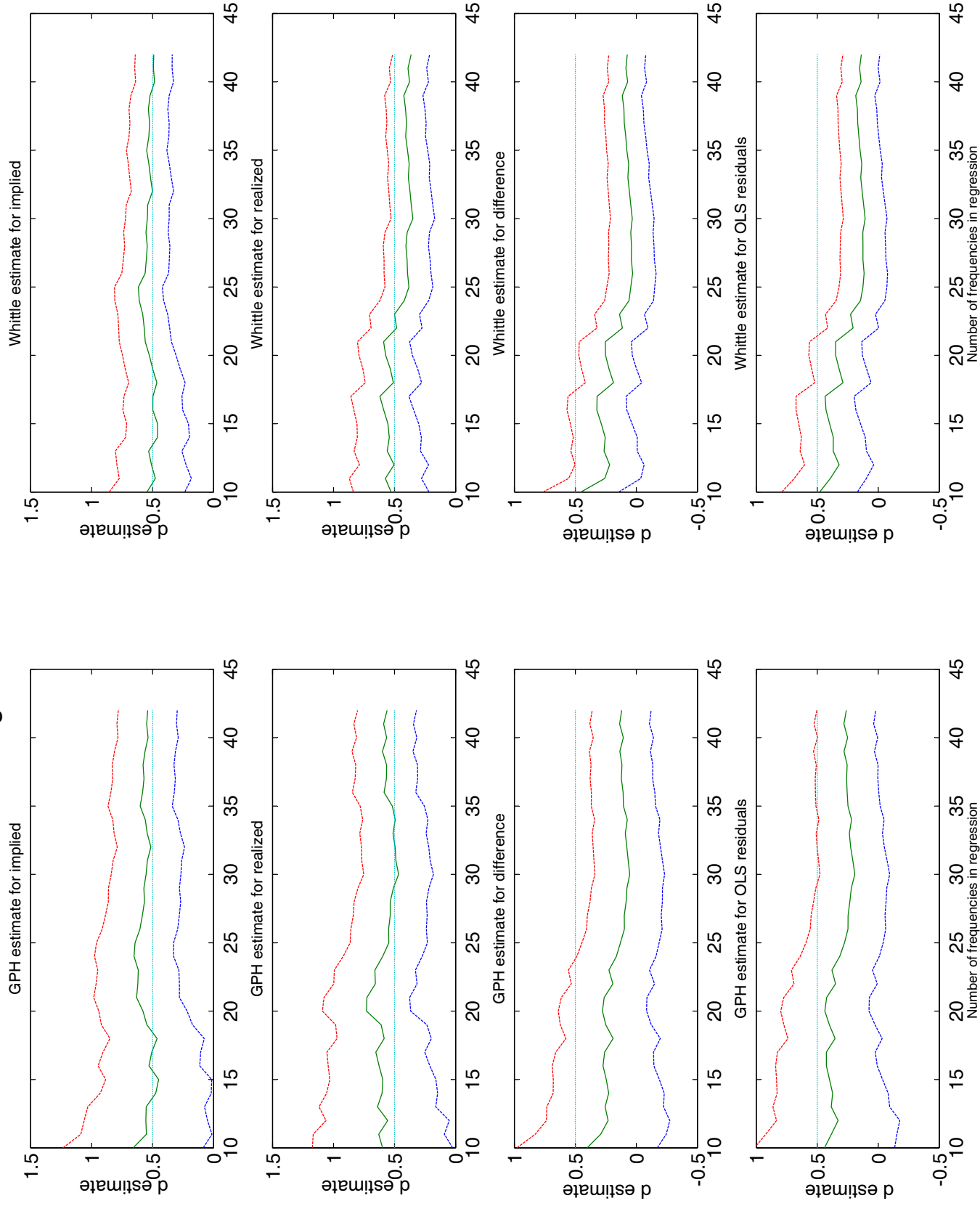


Figure 4. Narrow-band Estimates of Constant and Slope

