Long memory and the relation between implied and realized volatility*

Federico M. Bandi† Benoit Perron‡

This draft: January, 2003
First draft: April, 2001

Abstract

We argue that the conventional predictive regression between implied volatility (regressor) and realized volatility over the remaining life of the option (regressand) is likely to be a fractional cointegrating relation. Since cointegration is associated with long-run comovements, this finding modifies the usual interpretation of such regression as a study towards assessing option market efficiency (given a certain option pricing model) and/or short-term unbiasedness of implied volatility as a predictor for realized volatility, thereby rendering the conventional tests invalid.

We use spectral methods and exploit the long memory in the data to design an econometric methodology which is robust to the various issues that the literature on the relation between implied and realized volatility has proposed as plausible explanations for an estimated slope coefficient less than one, implying biasedness, in the standard predictive regression (measurement errors and presence of an unobservable time-varying risk premium, for instance).

Even though little can be said about market efficiency and/or short-term unbiasedness, which were the objects of the previous studies, our evidence in favor of a long-run one-to-one correspondence between implied and realized volatility series is rather strong. Simulation results confirm our findings.

*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 Tim Bollerslev, Silvia Gonçalves, Christian Gouriéroux, Javier Hidalgo, Joanna Jasiak, Chris Jones, Eric Renault, and participants at the 2002 Winter Meeting of the Econometric Society for helpful comments and 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 and the Fonds pour la Formation de Chercheurs et l'Aide à la Recherche (FCAR).

†*Graduate School of Business, The University of Chicago. Address: 1101 East, 58th street, Chicago, IL, 60637, USA. Tel. (773) 834-4352. E-mail: federico.bandi@gsb.uchicago.edu.*

‡*Université de Montréal and CRDE. Département de sciences économiques. Bureau C-6022. Pavillon Lionel-Groulx. 3150 Jean-Brillant. Montréal, QC, H3T 1N8, 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 realized volatility of the price of the asset an option is written on. As a consequence, in recent times there has been a substantial interest in verifying whether implied volatility is truly an unbiased forecast of future realized volatility: are the slope and the intercept of the regression of realized on implied volatility equal to 1 and 0, respectively (i.e., unbiasedness hypothesis)? Furthermore, is implied volatility an efficient estimate of the future realized volatility in the sense that economic variables belonging to the same information set as the former do not contribute predicting the latter (i.e., informational efficiency)?

The early studies on this topic generally maintained that the option implied volatility is an informative and biased forecast of future volatility. The evidence about informational efficiency is rather mixed (the reader is referred to Poteshman (2000) for a complete discussion of the extant literature). Recent work has pointed out that the evidence in favor of unbiasedness and efficiency is stronger once the three main issues that are believed to contaminate the relevant predictive regressions are properly accounted for, namely errors-in-variables (Christensen and Prabhala (1998), Poteshman (2000) and Chernov (2001), among others), overlapping data (Christensen and Prabhala (1998)) and missing terms (Poteshman (2000) and Chernov (2001)). A couple of examples are in order. Implied volatility is generally believed to be measured with an error that induces correlation between the regression residuals and the regressor, i.e., the implied volatility itself (Christensen and Prabhala (1998) and Poteshman (2000)). In addition, no-arbitrage pricing suggests the plausibility of a time-varying risk premium (Poteshman (2000) and Chernov (2001)) which is likely to covary with implied volatility and bring about non-orthogonality between residuals and regressors when omitted from the relevant predictive regression. Hence, missing terms (such as a time-varying risk premium) and errors-in-variables have the common feature of inducing correlation between regressor and residuals in the relevant regression, thereby biasing the least squares parameter estimates.

This work recognizes a feature of volatility series that has been discussed by many in the empirical literature but whose consequences have not been examined yet in predictive regressions of the type described in the preceding paragraphs: volatility is highly persistent.

The study of the implications of persistence for the conclusions drawn in the existing work on the predictive power of implied volatility represents the substantive core of the present paper.

Several authors have argued that spot and realized volatility may be well described by a long memory process (Ding et al. (1993), Baillie et al. (1996), Bollerslev and Mikkelsen (1996, 1999), and Andersen et al. (2001a, 2001b), among others). Other ways of modelling persistence in volatility, such as the components model of Engle and Lee (1999), the multi-factor specification of Chernov et al. (2001) and Alizadeh et al. (2002), or the jump models of Pan (2002), and Eraker et al. (2001), could also be entertained. Due to the well-known lack of consensus about the correct modelling approach and the necessity of taking a stance for the purpose of our subsequent analysis, we choose to adopt the specification that has received the most attention in applied work, that is we assume long memory.[1] While alternative specifications would require tight parametric specifications for the underlying spot volatility and, as a consequence, complications that we wish to avoid for the sake of parsimony of the model, long memory is now the standard way to interpret the persistence properties of nonparametric specifications for realized volatility series as in French et al. (1987) and Schwert (1989), as discussed in recent work by Andersen et al. (2001a, 2001b).

In addition to the long-range dependence properties of realized volatility, our results indicate existence of a fractional cointegrating relation between realized and implied volatility series. In accordance with recent developments in the literature on the forward discount anomaly (see Barnhart and Szakmary (1991), for example), cointegration necessarily points us towards a fundamental reinterpretation of the economic significance of the conventional predictive regression. In effect, such regression should be regarded as a characterization of the long-run equilibrium relation between volatility series rather than as a formal test of option market efficiency and/or short-term unbiasedness as in the extant literature on the subject.

Having made this point we proceed to the estimation of the model. The likely existence of a cointegrating relation (which could be either in the stationary or in the nonstationary

[1] Taylor (1986), Ding et al. (1993) and Dacorogna et al. (1993) are important early references on the relevance of long range dependence in volatility. Taylor (1986) finds persistence of the fractional type in absolute stock returns. Ding et al. (1993) and Dacorogna et al. (1993) note the same fact for the powers of daily returns and high frequency squared exchange returns, respectively. (The interested reader is referred to Baillie (1996) for a survey of the early work.) Baillie at al. (1996), Bollerslev and Mikkelsen (1996,1999), Comte and Renault (1998), Ebens (1999), Ray and Tsay (2000), Brunetti and Gilbert (2000), Li (2000) and Andersen et al. (2001a, 2001b), among others, provide more recent evidence.

region) between implied and realized volatility measures has to be taken into consideration when conducting inference since the standard framework for doing hypothesis testing does not apply. Nonetheless, fractional cointegration provides us with a natural set-up to tackle some of the problems that predictive regressions of this kind encounter. By virtue of the existence of cointegration between realized and implied volatility, we argue that contaminations inducing dependence between the regressor and the residuals (a time-varying risk premium, for instance) are likely to be of lower order of (fractional) integration than the regressor. In consequence, we use narrow band least squares spectral methods to estimate the model consistently *even* in the likely presence of stationary residuals correlated with a stationary regressor. As said, this is a unique feature of our methodology which crucially hinges on the statistical properties of the relation between realized and implied volatility series.

Let us be more clear. When assuming fractional cointegration in the nonstationary region, consistency of the least squares estimates in the presence of residuals correlated with the regressors is a non-surprising result and reflects well-known facts about conventional cointegrating relations of the I(1)/I(0) type, where the eventual correlation between errors and regressors only determines second error adjustments in the asymptotic distribution of the least squares parameter estimates (see Park and Phillips (1988), for example). In nonstationary fractionally cointegrating models, the least squares estimates converge at atypical rates and have nonstandard limiting distributions (see Kim and Phillips (1999b) and Robinson and Marinucci (2001), RM henceforth, for instance), thus requiring adjustments to the conventional standard errors for the purpose of reliable statistical testing. In the same models, narrow band least squares generally determine faster rates than standard least squares (RM (2001)). This observation provides a valid justification for using spectral methods based on a degenerating band of frequencies at the origin even in situations where more conventional methods, such as ordinary least squares, deliver consistent estimates. More importantly, in the stationary fractional cointegrating case the least squares estimates are known to be inconsistent but, again, narrow band spectral methods can be seen as a natural approach to guarantee consistency as first suggested by Robinson (1994a,b) and rigorously shown below in our framework. A simple discussion about one potential source of correlation between residuals and regressor, i.e., existence of a time-varying risk premium, will suffice here to clarify the intuition underlying this last important result. If the risk premium, which potentially enters the regression residuals, is of lower order of integration than implied volatility (the regressor), then its spectrum is dominated by that of implied volatility near zero frequency. Performing least squares over a set of frequencies in the vicinity of zero allows us to exploit this property and obtain consistent parameter estimates. In this work we are agnostic about the nature of the potential risk premium since its functional form is unknown (asset pricing theory does not give us any reliable indication) and its inclusion might determine likely mispecifications. More generally, we avoid being specific about the features of the non-orthogonality between residuals and regressor and find that, while option market efficiency and/or short-term unbiasedness cannot be tested confidently due to the properties of the data, evidence in favor of a long-term notion of the conventional unbiasedness hypothesis is rather strong.

Aside from the evident empirical appeal, modelling volatility series as long-memory processes finds an interesting theoretical justification in the work of Comte and Renault (1996, 1998), CR henceforth, which, in line with the prescriptions of the data, ties together the long memory properties of realized, implied and spot volatility. CR (1996, 1998) study long range dependence in continuous-time stochastic volatility models of option pricing. They show that if the underlying unobservable (spot) volatility series displays long-range dependence, then so do realized and implied volatility. Furthermore, the CR (1998) model provides a natural justification for testing the existence of a linear relationship between implied and realized volatility by virtue of the fact that the implied Black and Scholes volatility should be a proxy for the expected (under the equivalent martingale measure) realized volatility of at-the-money, short-term, options. CR (1998) show that, in addition to being an empirical fact, long range dependence in stochastic volatility models provides a rationale for the so-called "smile" (the U-shaped structure of implied volatilities across different strike prices). In particular, while standard stochastic volatility models capture the fact that the smile decreases as the time to maturity increases, such effect is more pronounced than in the data. Volatility persistence provides a valid justification for this effect. Interestingly, long range dependence in stochastic volatility is consistent with continuous-time no-arbitrage pricing. In particular, even though volatility is not a semimartingale in itself (since fractional Brownian motion is not a martingale, see Rogers (1997)), prices are (since volatility is not assumed to be a traded asset) and therefore admit an equivalent martingale measure (indeed, more than one since markets are typically not complete).

One last observation is in order. This paper discusses an instance where the persistence of volatility measures, which is a stylized fact based on our results and a good deal of 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 most volatility measures should be given more care any time a notion of volatility is employed as a dependent or independent variable in forecasting, as sometimes the case in the empirical finance literature. Any approach used must be able to capture persistence, and this is what our focus on a simple long-memory model accomplishes.

The paper is organized as follows. Section 2 contains a description of the data and tests for fractional cointegration of realized and implied volatility series. Section 3 briefly discusses an option pricing model that implies long-memory in realized and implied volatility given long memory in spot volatility. The same model provides us with economic restrictions to be tested. In Section 4 we describe the econometric methodology and show consistency of the parameter estimates. Section 5 contains simulations. The empirical results are laid out in Section 6. Section 7 concludes. Proofs and technicalities are in Section 8.

2 The data

We consider daily data on the VIX (i.e., the CBOE Market Volatility index) and the S&P 100 (OEX) index from January 1988 to May 2000.

The VIX is the implied volatility of a nontraded (synthetic) at-the-money option contract with one month to maturity. It is believed not to be affected by the problems that pollute standard implied volatility measures extracted from OEX contracts as in Christensen and Prabhala (1998), 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. The VIX was first used in predictive regressions of the type analyzed here by Chernov (2001). We refer the reader to Chernov's paper and to Whaley (2000) for additional details.

It is noted that Chernov employs daily overlapping data. We use monthly[2] non-overlapping observations to control the correlation structure of the errors as suggested by Christensen and Prabhala (1998) by taking the closing value of each month. We also multiply the VIX data by a constant factor equal to $\left(\frac{252}{365}\right)^{1/2}$ as in Schwert (2002) to account for the difference between trading and calendar days. As pioneered by French et al. (1987) and Schwert (1989), we compute the realized volatility over the remaining 30 days (one month) of the option as

$$\left(\sigma_{t,30}^{R}\right)^{2} = \frac{1}{30}\sum_{j=1}^{30} r_{t+j}^{2}, \tag{1}$$

where $r_t = \log\left(S_t/S_{t-1}\right)$ and S_t is the daily S&P 100 index at t.[3] Finally, we take the square root of both volatility measures.

To summarize, the data transformations result in two monthly time series, say σ_t^{BS} and σ_t^{R}, where σ_t^{BS} is the annualized (assuming 252 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 σ_t^{R} is the realized standard deviation of the S&P 100 index over the remaining life of the synthetic option (one month), respectively. Both series contain 152 non-overlapping observations.

Table 1 provides descriptive statistics for the two volatility measures, as well as for the least squares residuals from the regression

$$\sigma_t^{R} = \alpha + \beta\sigma_t^{BS} + u_t, \tag{2}$$

and the residuals assuming $\alpha = 0$ and $\beta = 1$. We notice that both implied and realized volatility display heavy tails and positive skewness. As typically the case with volatility measures (see Andersen et al. (2001b), for instance), a simple log transformations (not reported here) would almost restore normality. The mean of the implied volatility is larger than the mean of the realized volatility, suggesting a possible risk premium.

We now analyze the persistence properties of the data.

[2] >From a theoretical standpoint, the use of daily data as in Chernov (2001) does not affect the long memory properties of the series (which we verify below) since the class of fractionally integrated processes is self-similar implying that the degree of fractional integration is invariant to the sampling frequency (see Beran (1994), for example).

[3] French et al. (1987), Schwert (1989), and Christensen and Prabhala (1998) remove the mean of returns over the remaining life of the option before computing $\sigma_{t,30}^{R}$. Below we show that the estimator we use is more coherent with the interpretation of $\sigma_{t,30}^{R}$ as an estimate of the square root of the quadratic variation of the log-price process.

2.1 Long memory and fractional cointegration

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 consider both the standard Geweke and Porter-Hudak (GPH, henceforth) estimator (see GPH (1983)) and the Andrews and Guggenberger (AG, henceforth) estimator (see AG (2000)). Both are obtained as the least squares estimate of d in the frequency-domain regression

$$\ln I(\lambda_i) = \gamma_0 - d \ln\left(4\left(\sin^2\left(\frac{\lambda_i}{2}\right)\right)\right) + \sum_{j=1}^{J} \gamma_j \lambda_i^{2j} + \varepsilon(\lambda_i), \tag{3}$$

where $I(.)$ is the periodogram of the data computed at the harmonic frequencies $\lambda_i = \frac{2\pi i}{n}$ with $i = 1, ..., m < n$. The bandwidth parameter m is allowed to vary between $[\sqrt{n}]$ and $\left[n^{0.8}\right] + 1$, where $[x]$ is the integer part of x.

The GPH estimator is obtained with $J = 0$, while the AG estimator allows $J > 0$ (we set $J = 1$ in what follows). The asymptotic standard errors are derived from the result

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

where c_J is a constant that depends on the number of even powers of harmonic frequencies in the log-periodogram regression. For $J = 0$ (the GPH case), c_J is equal to 1, while it is 2.25 for $J = 1$ (see AG (2000) for other values of c_J and its formal definition). It is noted that adding the term λ_i^2 more than doubles the variance of the d estimates. Nonetheless, we report the AG estimator here since the bias improvement that this addition determines is sufficient to reduce the asymptotic mean squared errors of the d estimates relative to the GPH estimator.

The limiting distribution in (4) was obtained by AG (2000) under the assumption of stationarity ($d < \frac{1}{2}$). Nonetheless, it is known that when $J = 0$ (in the standard GPH case), such result is true (with $c_0 = 1$) both in the presence of stationary data (Robinson (1995)) and in the presence of nonstationary data with $\frac{1}{2} \leq d < \frac{3}{4}$ (Velasco (1999)). Additionally, for $\frac{1}{2} < d < 1$ the GPH estimator is known to be consistent (Velasco (1999)).[4]

In Figure 1 we report the d estimates and the corresponding 95% confidence bands for the original volatility series σ_t^{BS} and σ_t^R as well as for their difference (viz., the residuals obtained

[4] The literature cited previously assumes gaussianity. Nonetheless, milder conditions can be invoked for the results to be valid as shown by Kim and Phillips (1999).

by imposing a (1,-1) cointegrating vector) and the least squares residuals.[5]

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. Qualitatively, we find that this the case. As typical in fractional models, the d parameter estimates are fairly imprecise with large standard error bands (see Baillie (1996)); as expected, this is especially true for the AG estimator. Nonetheless, the d estimates are rather stable over a wide range of bandwidths and statistically indistinguishable across volatility measures. When combined with the lower degree of integration of both sets of residuals, this finding, which is robust to alternative bandwidth choices, suggests cointegration of the fractional type. Interestingly, cointegration in the nonstationary region appears to be plausible.[6]

Given these results, a simple model for the data might be

$$\sigma_t^R = \alpha + \beta\sigma_t^{BS} + u_t, \tag{5}$$

where $\sigma_t^R \sim$I(d_1), $\sigma_t^{BS} \sim$I(d_1) and $u_t \sim$I(d_2) with $d_2 < d_1$. Furthermore, we could assume that either $d_1 > 0.5$ with $d_1 + d_2 < 1$, implying nonstationarity and mild cointegration, or $d_1 < 0.5$, implying cointegration in the stationary region. Abstracting from the long memory properties of the data, regression (5), which derives (in an approximate form) from most option pricing models, has been employed in much existing work to test for option market efficiency and/or short-term unbiasedness of implied volatility as a predictor for realized volatility. Should we believe that fractional cointegration is a robust feature of the relation between implied and realized volatility series, then such tests appear hard to justify.

Naturally, we verify the robustness of our finding of fractional cointegration by applying the previous methodology to two additional data sets that were used in previous studies. The first data set 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.[7] The data is obtained

[5] Based on our previous discussion, ordinary least squares regressions are, in general, consistent only in the nonstationary case ($d \geq 0.5$). Since there is uncertainty as to whether the data is stationary or not (see below), the least squares results should be seen as illustrative only.

[6] CR (1998) find a fractional parameter equal to 0.67034 for the implied volatility (σ_t^{BS}) of options on the CAC40 of the Paris Stock Exchange.

[7] The use of post-crash (from January 1988) data that we make here (and made earlier) is due to the observation that markets were substantially less liquid in the pre-crash period (Poteshman (2001)). Jackwerth

by examining at-the-money call options with a month to expiration sampled right after the previous expiration date. This data is used in Christensen and Prabhala (1998). We refer the interested reader to their work for a detailed description. The second data set consists of monthly data on options on Deuschmark-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 Jorion's sample covers the January 1985 to February 1992 period. 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. It is noted that this market moved from quarterly option expiration dates to monthly expiration dates in 1987. Here, we work with data from 1990 to avoid any potential adjustment/learning effects that this change might have caused.

The results are presented in Figures 2 and 3. We confirm that fractional cointegration appears to be a stylized fact, thereby triggering a fundamentally different interpretation of the economic significance of regression (5).

As a final robustness check (not reported here for brevity[8]), we consider two transformations of the data that have been implemented in previous research, namely we square the volatility series and take logarithms.[9] The conclusions we draw are the same.

Having pointed out that fractional cointegration necessarily changes our understanding and interpretation of a well-known economic relation, we now turn to a model that justifies estimating a regression equation like (5) above.

A final caveat is in order. In our preliminary investigation of the data we use graphical methods to assess the presence of cointegration and evaluate its nature. There are two reasons

and Rubinstein (1996), for example, report structural changes in the S&P 500 (SPX) market following the October 1987 crash. It is apparent that such changes are likely to have affected the OEX market as well. It should be noted, though, that our main finding, i.e., the existence of a cointegrating relation between volatility series, is robust to the inclusion of the pre-crash data. The full data simply displays fatter tails than reported in Table 1 and less persistence, thus inducing cointegrating relations that are closer to, or deeper into, the stationary region.

[8]Corresponding tables and figures can be provided by the authors upon request.

[9]Our original transformation, namely $\sigma^{R,BS}$, 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 (2001) use $\left(\sigma^{R,BS}\right)^2$. The log transformation is in Christensen and Prabhala (1998).

for this. First, we show robustness of the results to the choice of the number of frequencies. Second, by looking at the residuals obtained from imposing the cointegrating vector $(1, -1)$ as well as the least squares residuals, which are admittedly biased in the stationary case (see our previous discussion), we provide rather sharp evidence (in favor of cointegration) that allows us to by-pass the problems that would be posed by the implementation of the existing tests. For instance, the choice of the bandwidth when estimating consistently the cointegrating vector by narrow band spectral methods and the distribution of the residuals of the cointegrating vector are quite problematic (see Brunetti and Gilbert (2000) for a recent discussion).

3 Long-memory in implied volatility: a theoretical justification

Most option pricing models we are aware of imply an approximate linear relation between implied volatility and realized volatility over the life of the option. This observation provides justification for regressing realized volatility onto implied volatility in a linear fashion since the appeal of the maintained hypothesis of the existence of a linear relation between the two volatility measures appears to be largely unaffected by the validity of a specific option pricing model.

Here we use the set-up in CR (1998) to introduce formally fractionally integrated volatility in continuous-time option pricing and derive a meaningful testable framework. Coherently with the short-range dependence case discussed by others (see Poteshman (2000) for a review), this model implies an approximate linear relation between the two volatility measures that are the object of the present study. As we discuss below, such relation is robust to various specifications of the underlying fundamentals.

Assume a stochastic volatility model with long range dependence of the type[10]

$$\frac{dS_t}{S_t} = \mu_S(t, S_t)dt + \sigma_t dW_t^1 \tag{6}$$

$$d(\ln \sigma_t) = k(\theta - \ln \sigma_t)dt + \gamma dW_{dt}^2, \tag{7}$$

where

[10] The interested reader is referred to CR (1996) for an introduction to long memory continuous-time models and Comte (1996) for a discussion of simulation and estimation methods for the same class of models.

$$W_{dt}^2 = \frac{1}{\Gamma(1+d)} \int_0^t (t-s)^d dW_s^2 \qquad 0 < d < \frac{1}{2} \tag{8}$$

is a version[11] of fractional Brownian motion and $\{W^1(t), W^2(t)\}$ is a standard Wiener process on the plane. The log volatility process displays long range dependence since the spectrum

$$f_{\ln\sigma_t \ln\sigma_t}(\lambda) = \frac{\gamma^2}{\Gamma^2(1+d)\lambda^{2d}} \frac{1}{\lambda^2 + k^2} \tag{9}$$

is unbounded at the origin.

A European call option with strike price K and maturity T written on the financial asset with price S_t has a value at time t which is given by the expectation of the Black and Scholes model, where the expectation is taken with respect to the equivalent martingale measure Q and the implied Black and Scholes volatility is replaced by the continuous time average of σ_t^2 over the period of interest (CR (1998)). Specifically, the price C_t of the option can be written as

$$C_t = \Im_t^Q \left(C^{BS} \left(S_t, \sqrt{U_{t,T}}, K, T-t \right) \right), \tag{10}$$

where

$$U_{t,T} = \frac{1}{T-t} \int_t^T \sigma_s^2 ds = \frac{1}{T-t} \left[\log S\right]_{t,T}, \tag{11}$$

[11] MR (1999) define a rescaled version of it, namely

$$\overline{W}_d^2(t) = (2d+1)\Gamma(1+d) W_d^2(t) \qquad t \geq 0$$

as "type II fractional Brownian motion". The main difference between type II and what MR (1999) call type I (see Taqqu (1979) and Samorodnitsky and Taqqu (1994)), i.e.

$$\overline{B}_d^2(t) = \frac{1}{A(d)} \int_{-\infty}^0 \left\{ (t-s)^d - (-s)^d \right\} dW(s) + \frac{1}{A(d)} \int_0^t (t-s)^d dW(s)$$

for $t \geq 0$ with

$$A(d) = \left\{ \frac{1}{2d+1} + D(d) \right\}^{1/2}, \; D(d) = \int_0^\infty \left\{ (1+s)^d - s^d \right\}^2 ds$$

lies in the nonstationarity of the increments in the case of type II. It readily appears that $\overline{W}_d^2(t)$ is simply a truncated version of $\overline{B}_d^2(t)$. Just like $\overline{B}_d^2(t)$ can be obtained as the weak limit of a (rescaled) sum of stationary linear processes displaying long-range dependence (see Davydov (1970), for example), $\overline{W}_d^2(t)$ can be the weak limit of a (rescaled) sum of truncated (asymptotically stationary) linear processes with long-range dependence (see Akonom and Gourieroux (1987), Silveira (1991) and MR (2000)).

$[\log S]_{t,T}$ is the quadratic variation of the log-price process between t and T and $\Im_t^Q(.) = \mathbf{E}^Q(.|\Im_t)$. Hence, the following proposition readily follows.

Proposition 1. *Given an underlying asset whose price dynamic is driven by (6), (7) and (8), the value of an at-the-money, short-term, European option is such that the Black and Scholes implied standard deviation is approximately equal to the expected (under the objective measure P) square root of the averaged quadratic variation of the log price process plus a time-varying risk-premium, i.e.,*

$$\sigma_t^{BS} \approx \Im_t^P \left(\sqrt{U_{t,T}} \right) + RP_t. \tag{12}$$

Assuming that the time-varying risk premium RP_t can be quantified, expression (12) implies the testable restrictions $\alpha = 0$, $\beta = 1$ and $\gamma = -1$ on the regression model

$$\sqrt{U_{t,T}} = \alpha + \beta \sigma_t^{BS} + \gamma RP_t + \varepsilon_{t,T}, \tag{13}$$

where $\varepsilon_{t,T}$ is a forecast error satisfying $\Im_t^P(\varepsilon_{t,T}) = 0$. Unfortunately, simple no-arbitrage principles do not provide us with an expression for RP_t. 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 a similar exercise in a stochastic volatility model with short-range dependence). From an econometric standpoint, any assumption about RP_t is, therefore, arbitrary and potentially very misleading. In the sequel we will treat this term as unobservable. In other words, we will account for its presence but will not impose a tight structure on it to avoid likely mispecifications. We will come back to this issue.

Some additional observations are in order. First, it is noted that the pricing formula that justifies an approximate linear relation between implied and realized volatility, viz., (12) above, simply derives from the Q distribution of the price process being conditionally (on information and the volatility path) lognormal. In fact, the model is generalizable to more involved volatility structures provided the drift and the instantaneous volatility of the (positive-valued) volatility process only depend on the underlying volatility itself. Furthermore, the adapted drift process $\mu_S(t, S_t)$ is completely unrestricted.

Second, as shown by CR (1998), the Black and Scholes implied volatility displays long memory in this framework (see their Proposition 4.2). This observation implies that quadratic variation should also display long memory (since its order of integration is equal to the maximum order of integration of the three terms on the right-hand side of (13)). Hence, consistently with our findings in Section 2, the model presented in this section suggests a linear relation between two variables displaying long-memory characteristics. Should $\gamma RP_t + \varepsilon_{t,T}$ be of lower order of (fractional) integration than both volatility measures, as argued in Section 2 from an empirical perspective, then (13) constitutes a (fractional) cointegrating relation.

While the CR (1998) model is a useful framework to derive testable restrictions, as in the short-range dependence case covered elsewhere, it is apparent that the empirical appeal of a linear relation between realized and implied volatility series goes beyond the strict validity of the CR (1998) set-up.

We now turn to the estimation of model (13).

4 The econometric model

Four main econometric issues have been discussed in the literature on the predictive power of implied volatility.

First, the quadratic variation of the log price process is measured with error (see Poteshman (2000) and Chernov (2001), for example). What is generally called "realized volatility" is in fact an estimate of the random quadratic variation process of the log-price process. Such estimate is known to be consistent in probability provided the distance between observations goes to zero (see Protter (1990), inter alia). As a consequence, observations that are sampled at high frequencies provide a natural tool to estimate $[\log S]_{t,T}$. Andersen et al. (2001, 2001b) discuss the general issue of quadratic variation estimation using high-frequency data. Poteshman (2000) and Chernov (2001) apply this idea to predictive regressions of the kind analyzed in this work.

Second, the implied volatility series is also subject to measurement error (Christensen and Prabhala (1998), for example, mention the potential nonsynchronous measurement of options and index levels, early exercise and dividends, bid-ask spreads as well as the wild card option in the S&P100 (OEX) market).

Third, as stressed earlier, the time-varying risk premium is unobservable. Unless we are eager to be more specific about the features of the model, neglecting it implies potentially inconsistent least squares estimates of the parameter β since the residuals are likely to be correlated with the regressor, viz., σ_t^{BS} (Using the Heston (1993) model, for example, Chernov (2001) writes RP_t as a linear function of the unobserved stochastic volatility σ_t^2 but σ_t^2 is trivially correlated with σ_t^{BS}).

Fourth, the use of overlapping data as in Canina and Figlewski (1993) and Chernov (2001) induces a correlation structure in the errors which biases downward the least squares standard errors (see Christensen and Prabhala (1998) and Jorion (1995), for example).

In general, the second and the third issue have the common feature of causing inconsistent least squares parameter estimates when considering conventional stationary models.

Here, we accommodate both issues, as well as an imprecisely measured quadratic variation process, by explicitly modeling an error structure that is coherent with measurement errors and missing terms in the context of a cointegrating relation for implied and realized volatility series. We neglect the last issue, viz., correlated errors induced by overlapping data, since we employ non-overlapping observations in this study (see Section 2). Nonetheless, our framework would be robust to the use of overlapping data. Define

$$\sqrt{\frac{1}{T-t}\left[\log S\right]_{t,T}} + u_{t,T}^{R} = y_{t,T}, \tag{14}$$

$$\sigma_t^{BS} + u_t^{BS} = x_t, \tag{15}$$

$$-\left\{RP_t - \mathbf{E}\left(RP_t\right)\right\} = \delta_t, \tag{16}$$

$$-\mathbf{E}\left(RP_t\right) = \alpha, \tag{17}$$

where $y_{t,T}$ and x_t are the observable volatility series (i.e., realized and implied, respectively). Then, (13) suggests the testable model

$$y_{t,T} = \alpha + \beta\left(x_t - u_t^{BS}\right) + \delta_t + u_{t,T}^{R} + \varepsilon_{t,T}, \tag{18}$$

which implies

$$y_{t,T} = \alpha + \beta x_t - \beta u_t^{BS} + \delta_t + u_{t,T}^R + \varepsilon_{t,T} \tag{19}$$

and

$$y_{t,T} = \alpha + \beta x_t + \varepsilon_{t,T}^* \tag{20}$$

with $\varepsilon_{t,T}^* = -\beta u_t^{BS} + \delta_t + u_{t,T}^R + \varepsilon_{t,T}$, where the measurement errors u_t^{BS} and $u_{t,T}^R$ and the forecast error $\varepsilon_{t,T}$ are serially independent shocks with short-range dependence. Consistently with Robinson (1994), we model x_t and $\varepsilon_{t,T}^*$ as linear processes with spectra satisfying

$$f_1(\lambda) \approx L_1\left(\frac{1}{\lambda}\right)\lambda^{-2d_1} \quad \text{as} \quad \lambda \to 0^+, \tag{21}$$

$$f_2(\lambda) \approx L_2\left(\frac{1}{\lambda}\right)\lambda^{-2d_2} \quad \text{as} \quad \lambda \to 0^+, \tag{22}$$

respectively, where $L(\lambda)$ is a slowly varying function at infinity, i.e., a positive, measurable function so that

$$\frac{L(t\lambda)}{L(\lambda)} \to 1 \quad \text{as} \quad \lambda \to \infty, \quad \forall t > 0. \tag{23}$$

We recognize the existing fractional cointegration between $y_{t,T}$ and x_t by assuming that the parameters d_1 and d_2 conform to the requirement $d_1 > d_2$ with $0 \leq d_2 < \frac{1}{2}$. Based on our discussion of the data in Section 2, the fractional parameter d_1 can be either in the stationary region (i.e., $0 < d_1 < \frac{1}{2}$) or in the nonstationary region (i.e., $d_1 \geq \frac{1}{2}$).

Clearly, d_2 might be equal to zero, implying short-range dependence. The possible fractional integration of the aggregate error term $\varepsilon_{t,T}^*$ is necessarily induced by the missing (de-meaned) risk-premium δ_t since the shocks u_t^{BS} and $u_{t,T}^R$ and the forecast error $\varepsilon_{t,T}$ display short range dependence. The potential long memory properties of the risk premium can be understood easily. For instance, a large amount of literature writes the unobservable risk premium as a function (which is often linear, see Chernov (2001), for example) of the underlying spot volatility. But the underlying unobservable volatility is long-range dependent and, in general, transformations of long-range dependent series are long-range dependent (see Dittman and Granger (2000), for instance).

Additionally, we necessarily allow for potential correlation between the risk premium δ_t and the implied volatility proxy x_t (and, as a consequence, between $\varepsilon^*_{t,T}$ and x_t). Again, the plausibility of this feature of the model can be gauged easily. Trivially, if the unobserved risk premium depends on the underlying spot volatility series (as implied by many existing models), then the correlation between the implied volatility proxy (the regressor) and the regression residuals follows immediately. Chernov (2001) provides a discussion of the importance of this feature of the model in determining inconsistent least squares estimates in the stationary, short-range dependent, framework that he considers.

Coherently with our use of non-overlapping data, we write $T = t + 1$. Thus, the regression equation becomes

$$y_{t+1} = \alpha + \beta x_t + \varepsilon^*_{t+1}, \tag{24}$$

with ε^*_{t+1} defined as above. It is apparent that the error term ε^*_{t+1} is measurable with respect to $\Im^P_{t+1}$. Its conditional first moment cannot be zero since u^{BS}_t and δ_t are $\Im^P_t$-measurable. Hence, $\Im^P_t(\varepsilon^*_{t+1}) = -\beta u^{BS}_t + \delta_t$. Nevertheless, $\mathbf{E}(\varepsilon^*_{t+1}) = 0$.

Model (12) suggests that β should be equal to one and α should be equal to the expected risk premium (with a negative sign), i.e., $-\mathbf{E}(RP_t)$. We wish to test whether $\alpha = 0$ and $\beta = 1$. It should be pointed out that, contrary to the existing literature, an estimated α which is significantly different from zero cannot be invoked to rule out "unbiasedness." In fact, as shown earlier, in the presence of a non-vanishing expected risk premium the true α should be different from zero. Of course, a zero α does not imply that volatility risk is not priced. In fact, it might be priced and be time-varying but have a zero unconditional first moment. As a consequence, a rejection of the hypothesis $\alpha = 0$ would be evidence in favor of the existence of a priced premium. Should we fail to reject the hypothesis $\alpha = 0$, then we could only conclude that on average the volatility risk is not priced (see Pham and Touzi (1996)). Nonetheless, this result would provide some empirical evidence against studies that have pointed out the existence of a significantly negative premium (Benzoni (1998) and Pan (2000), inter alia). Interestingly, our framework potentially allows us to test for the existence of a priced volatility premium (on average) without making assumptions on the functional form of its relationship with the underlying volatility series.

We now turn to the details of the econometric procedure.

4.1 Predictive spectral regression

We notice that

$$y_{t+1} = \alpha + \beta x_t + \varepsilon^*_{t+1} \tag{25}$$

implies

$$y_{t+1} - \overline{y} = \beta \left(x_t - \overline{x}\right) + \left(\varepsilon^*_{t+1} - \overline{\varepsilon}^*\right), \tag{26}$$

where $\overline{b} = \left(\sum_{t=0}^{n-1} b_t\right)/n$. We define the discrete Fourier transforms,

$$\varpi_y(\lambda_s) \quad : \quad = \frac{1}{\sqrt{2\pi n}} \sum_{t=0}^{n-1} y_{t+1} e^{it\lambda_s} = \frac{1}{\sqrt{2\pi n}} \sum_{t=0}^{n-1} \left(y_{t+1} - \overline{y}\right) e^{it\lambda_s} \tag{27}$$

$$\varpi_x(\lambda_s) \quad : \quad = \frac{1}{\sqrt{2\pi n}} \sum_{t=0}^{n-1} x_t e^{it\lambda_s} = \frac{1}{\sqrt{2\pi n}} \sum_{t=0}^{n-1} \left(x_t - \overline{x}\right) e^{it\lambda_s} \tag{28}$$

$$\varpi_{\varepsilon^*}(\lambda_s) \quad : \quad = \frac{1}{\sqrt{2\pi n}} \sum_{t=0}^{n-1} \varepsilon^*_{t+1} e^{it\lambda_s} = \frac{1}{\sqrt{2\pi n}} \sum_{t=0}^{n-1} \left(\varepsilon^*_{t+1} - \overline{\varepsilon}^*\right) e^{it\lambda_s} \tag{29}$$

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

$$I_{yx}(\lambda_s) = \varpi_y(\lambda_s)\overline{\varpi_x(\lambda_s)}. \tag{30}$$

where $\overline{\varpi_x(\lambda_s)}$ is the complex conjugate of $\varpi_x(\lambda_s)$. In consequence, the narrow band least squares (NBLS) estimate of β is defined as

$$\widehat{\beta}_{n,m} = \frac{\widehat{F}_{xy}(1,m)}{\widehat{F}_{xx}(1,m)} = \frac{\widehat{F}_{xy}(0,m) - \overline{xy}}{\widehat{F}_{xx}(0,m) - \overline{x}^2}, \tag{31}$$

where

$$\widehat{F}_{xy}(l,m) = \frac{2\pi}{n} \sum_{s=l}^{m} I_{yx}(\lambda_s) \qquad l = 1, 0 \tag{32}$$

with $m < \frac{n}{2}$. We estimate α using

$$\widehat{\alpha}_{n,m} = \sqrt{\frac{2\pi}{n}}\varpi_y(0) - \widehat{\beta}_{n,m}\left(\sqrt{\frac{2\pi}{n}}\varpi_x(0)\right). \tag{33}$$

Some remarks are in order. We start with the stationary case (i.e., $0 < d_1 < \frac{1}{2}$). As discussed by Robinson (1994) and mentioned in RM (2001), even though least squares estimation is inconsistent in the presence of likely correlation between the regressors and the errors, the NBLS estimate of the slope parameter in the same regression is generally consistent provided the regressors are linked by a cointegrating relation. In our framework, correlation between the regressors and the error term is induced by the likely correlation between x_t (proxy for implied volatility) and δ_t (the de-meaned unknown risk premium). Hence, fractional cointegration in the stationary case allows us to get around the problems posed by the existence of a time-varying risk premium possibly correlated with implied volatility. 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 zero frequency. In consequence, we perform least squares over a degenerating band of frequencies in the neighborhood of the origin to exploit this property. The intercept estimator $\widehat{\alpha}_{n,m}$ is trivially consistent under the assumptions made provided $\widehat{\beta}_{n,m}$ is consistently estimated.

Theorem 1. *Consider the predictive model for a realized volatility proxy $y_{t,T}$ given an implied volatility proxy x_t*

$$y_{t,T} = \alpha + \beta x_t + \varepsilon^*_{t,T} \tag{34}$$

with $\varepsilon^*_{t,T} = -\beta u_t^{BS} + \delta_t + u_{t,T}^R + \varepsilon_{t,T}$, where u_t^{BS} and $u_{t,T}^R$ are measurement errors, $\varepsilon_{t,T}$ is a forecast error and δ_t is a time-varying risk premium. Assume u_t^{BS} and δ_t are correlated with x_t. Also, assume the spectra of x_t and δ_t satisfy

$$f_1(\lambda) \approx L_1\left(\frac{1}{\lambda}\right)\lambda^{-2d_1} \quad \text{as} \quad \lambda \to 0^+, \tag{35}$$

$$f_2(\lambda) \approx L_2\left(\frac{1}{\lambda}\right)\lambda^{-2d_2} \quad \text{as} \quad \lambda \to 0^+, \tag{36}$$

respectively, where L_1 *and* L_2 *are regularly varying functions at infinity and* $d_1 > d_2$ *with* $0 \leq d_1, d_2 < \frac{1}{2}$. *Then, the NBLS estimates defined in (31) and (33) are consistent in probability as* $1/m + m/n \to 0$ *with* $n \to \infty$.

We now turn to the nonstationary case (i.e., $d_1 \geq \frac{1}{2}$). Typically, in cointegrating relations of the I(1)/I(0) type consistency is not an issue in the presence of residuals that are correlated with the regressors since endogeneity only requires second order adjustments in the asymptotic distribution of the least squares parameter estimates (Phillips and Park (1998)). The same result trivially emerges from fractionally cointegrating regressions of the type analyzed here as shown by Kim and Phillips (1999b) and RM (2001), inter alia. Nonetheless, NBLS methods generally entail faster rates of convergence than ordinary least squares (see RM (2001)), thus providing justification for their use even in situations where more conventional tools could employed.

In order to assess the statistical significance of the regression estimates, we rely on subsampling (see Politis, Romano, and Wolf (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 some (rather mild) 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 Bertail, Politis, and Romano (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 our framework, the conditions of Theorems 8.2.1 and 8.3.1 of Politis, Romano, and Wolf (1999) are satisfied under some very standard asymptotic requirements: the size b of the subsamples grows to infinity and the ratio b/n vanishes as the sample size diverges to infinity in both the stationary and in the nonstationary case. This implies that we can obtain consistent estimates of the convergence rates of the slope and constant estimators as well as valid confidence intervals without knowing whether the stationary or the nonstationary case apply to our case. To preserve the correlation between the regressors and the errors, we do the resampling on the couple (y, x) and choose subsamples of size $b = \left[n^{\frac{2}{3}}\right]$ where $[x]$ denotes, as usual, the integer part of x (see Politis and Romano (1994) and Bertail, Politis, and Romano (1999) for a similar choice). We compute statistics on all $n-b+1$ such subsamples of b successive observations and choose m_b, the number of Fourier frequencies to compute the narrow band estimator for each subsample, to be the same fraction of the sample size as in the original data, i.e., $\frac{m_b}{b} = \frac{m}{n}$.

In the next section we simulate an econometric model that is coherent with the assumptions of Theorem 1 by using the set-up that was proposed by CR (1998). We also evaluate the properties of the proposed estimators.

5 Simulations

We employ the CR (1998) long memory stochastic volatility model (c.f. (6) and (7) above) to analyze the impact of long memory on the predictive regression between implied and realized volatility as well as the behavior of the proposed narrow band least squares estimates. We use the discretized version of the model provided by Comte and Renault (1998) and the algorithm in Comte (1996) to generate sample paths. We compute option prices by numerical integration assuming that volatility risk is not compensated.

The discretized model is:

$$\ln S_{t+\Delta} = \ln S_t + \left(r - \frac{\sigma_t^2}{2}\right)\Delta + \sqrt{\Delta}\sigma_{t+\Delta}\varepsilon_{t+\Delta}^1 \tag{37}$$

$$\ln \sigma_t = \theta + \sum_{j=1}^{[t/\Delta]} \frac{(t - j\Delta)^d}{\Gamma(1+d)} \left[x_{j\Delta}^{(d)} - x_{(j-1)\Delta}^{(d)}\right] \tag{38}$$

$$x_{t+\Delta}^{(d)} = e^{-k\Delta} x_t^{(d)} + \gamma \left(\frac{e^{-2k\Delta} - 1}{-2k}\right)^{\frac{1}{2}} \varepsilon_{t+\Delta}^2 \tag{39}$$

$$\begin{pmatrix} \varepsilon_{t+\Delta}^1 \\ \varepsilon_{t+\Delta}^2 \end{pmatrix} \sim iidN\left(0, I_2\right) \tag{40}$$

We set $x_0^{(d)} = 0$ and normalize the log stock price to be 0 at time 0. This normalization is innocuous since we only look at at-the-money options. We take Δ to be half a day. This implies that each replication involves computing 152 option prices, each of which is inverted to obtain the BS implied volatilities. Each option price is computed from 20,000 simulations. Since we use antithetic variables, we have 10,000 different draws of the underlying random errors.

In order to reduce the effect of initialization, for each replication we generate a single series of $2 \times 22 \times 152 = 6,688$ realizations of the relevant continuous-time process plus 1,000 points to be discarded. Realized volatility is computed by summing the squares of the daily returns over the life of the option, i.e., 22 squared returns corresponding to every other simulated value of the log price process. The path of the stock price used to compute realized volatility is chosen at random among the 20,000 realizations. We replicate the procedure 1,000 times.

The parameter values are set at those estimated on our daily data on the S&P 100 index between January 1988 and July 2000. We estimate the model using the efficient method of moments (EMM) code provided by Gallant and Tauchen (see Gallant and Tauchen (1996, 2002) for details). The estimated parameter values are (in annual terms):

$$r = 0.00962 \quad (41)$$

$$\theta = -4.673 \quad (42)$$

$$k = 1.561 \quad (43)$$

$$\gamma = 0.388 \quad (44)$$

$$d = 0.506. \quad (45)$$

Finally, in order to introduce a likely source of correlation between implied volatility and regression residuals, we add a measurement error to the implied volatility series.[12] The error term is independent and identically distributed through time. Specifically, we assume that it is normally distributed with zero mean and a variance that is equal to the variance of implied volatility over the corresponding replication.

For each replication we run a linear regression of realized volatility on a constant and the implied volatility proxy. This regression is estimated by ordinary least squares and by

[12] We could have chosen a different source of correlation between regressor and regression residuals. For instance, coherently with our previous discussion, we might have assumed that volatility risk is compensated and set the risk premium as to induce correlation with implied volatility. While this alternative procedure would not have been more informative, it would also have been substantially more arbitrary due to the fact that little is known about credible functional forms for the unknown premium.

narrow band least squares with 4 choices of bandwidth corresponding to different powers of the sample size, namely $n^{0.5}$, $n^{0.6}$, $n^{0.7}$, and $n^{0.8}$. With the sample size that we are using (152 observations), these choices amount to including 12, 20, 33, and 55 Fourier frequencies in the relevant regression.

In Table 2 we report the means of the various quantities over the replications as well as the standard deviations (in parentheses, underneath). The first column of the table refers to the estimation of the constant, α, while the second column refers to the estimation of the slope, β. The first thing to notice is that ordinary least squares provides very biased results with a mean estimated slope of about 0.35. Moreover, the R^2 from this regression is barely above 5%. Narrow band least squares remove almost all of this bias with a mean slope of about 0.9 when using the smaller bandwidth. We also notice a downward trend to the estimates of β (and, of course, an upward trend to the estimates of α) as the number of Fourier frequencies in the spectral regression is increased since more and more short-run noise is added to the regression. This result is coherent with theory and intuition. It is also consistent with the empirical results that are reported in the subsequent section. As can be expected, the precision of the estimator decreases as we reduce the number of Fourier frequencies.

In Table 3 we report the results relating to the estimation of the degree of long memory using both the GPH and the AG estimator. Again, we consider 4 bandwidths corresponding to $n^{0.5}$, $n^{0.6}$, $n^{0.7}$, and $n^{0.8}$ with $n = 152$. As in the case of the slope estimates, we see a downward trend in the estimates of d as we increase the number of frequencies in the estimation of this parameter. The only exception to this rule is the AG estimator for implied volatility. Consistently with theory, the AG estimator is more variable with an increase in standard deviation of roughly a third for implied volatility and almost a half for realized volatility. While spectral methods are useful in removing the short-run noise that is associated with both volatility measures, estimation by spectral methods is subject to a standard bias-variance trade-off in the choice of bandwidth. Thus, care should be exercised when implementing and interpreting this approach.

6 Empirical results

Narrow band estimation of the predictive regression model (24) is presented in Figure 4 as a function of the number of Fourier frequencies included in the spectral regression. The first row of the figure corresponds to the results for the S&P 100 and VIX data, the second row to the results for the S&P 100 contracts, and the last row to the results for the DM options. Note that the last point on each graph (the point that includes all the Fourier frequencies) refers to the ordinary least squares estimate. The left panels contain results for the constant α, while the right panels provide results for the slope coefficient β. We also report 95% confidence bands for both sets of parameters based on the subsampling algorithm described in Section 4. Finally, a horizontal line indicates the location of 0 for the constant and 1 for the slope for ease of interpretation.

We first discuss the S&P 100 - VIX case. Consistently with intuition and the results in the previous section, there is a slight downward trend to the slope coefficients (such trend is more evident when analyzing the other two data sets). As the number of frequencies in the regression increases, a greater proportion of the 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 the intercept estimates are concerned, their values point to the existence of a slight (average) compensation for volatility risk.

A caveat is in order. While the data supports long-term unbiasedness in the sense that we cannot reject the null $\beta = 1$, there are at least two reasons why it is theoretically conceivable to estimate a slope coefficient different from one. The first one has a statistical flavor, while the second one has a superior economic appeal.

We start with the first explanation. Even if model (12) were the correct model, should realized volatility be cointegrated with implied volatility *and* the risk premium and the risk premium be cointegrated with implied volatility, then the cointegrating vector between realized and implied volatility would not necessarily be $(1, -1)$. Based on models (12) and (24), consider

$$a_{t+1} = y_{t+1} - x_t - \delta_t \tag{46}$$

and

$$b_t = \alpha x_t + \beta \delta_t \tag{47}$$

with a_{t+1} and b_t integrated of lower order than x_t (implied volatility proxy) and δ_t (risk premium). Then, it would follow

$$y_{t+1} = x_t + \left(\frac{b_t}{\beta} - \frac{\alpha}{\beta} x_t \right) + a_{t+1}, \tag{48}$$

which would be equivalent to

$$y_{t+1} = \left(1 - \frac{\alpha}{\beta} \right) x_t + \frac{b_t}{\beta} + a_{t+1}. \tag{49}$$

In other words, $\widehat{\alpha}_{n,m}$ and $\widehat{\beta}_{n,m}$ would be consistent estimates of the parameters of the true cointegrating vector (i.e., the true long-run relation between implied and realized volatility), but the slope would be $\left(1 - \frac{\alpha}{\beta}\right)$ even if model (12) were correct. Clearly, $\left(1 - \frac{\alpha}{\beta}\right)$ can be smaller or larger than 1 depending on the features of the cointegrating vector between implied volatility and the risk premium. Of course, even if such vector existed, it would be impossible to estimate α and β separately since the underlying risk premium is unobservable.

We now turn to the second explanation for an estimated slope different from one, namely the existence of a non-zero correlation between stock returns and the unobservable volatility process (the so-called leverage effect).[13] While the previous observation depends crucially on the fractional (co-)integration of the volatility series that are the objects of this study, the present explanation emerges readily even from more conventional settings where short-range dependence is satisfied. In order to stress that the result does not hinge on the long memory properties of the data and for the sake of simplicity, we use a conventional short-range dependence framework to illustrate this point. We largely follow Renault (2001). Consider the following model for the risk-neutral (Q) dynamics of the stock return process

$$\frac{dS_t}{S_t} = r_t dt + \rho \sigma_t dW_t^{2*} + \sqrt{1-\rho^2} \sigma_t dW_t^{1*} \tag{50}$$

$$d \ln \sigma_t = g^*(\sigma_t) dt + \sigma_t dW_t^{2*}. \tag{51}$$

[13] We thank Eric Renault for pointing this out to us.

As usual, W_t^{1*} and W_t^{2*} are independent driftless Brownian motions under Q. Notice that $Corr_t\left(\frac{dS_t}{S_t}, d\ln\sigma_t\right) = \rho dt$. Then, the following proposition applies.

Proposition 2. *Given an underlying asset whose risk neutral price dynamics is driven by (50) and (51), the value of an at-the-money, short-term European option is such that the implied Black and Scholes standard deviation is approximately equal to the square root of one minus the squared correlation between return and volatility process, i.e.,* $\sqrt{1-\rho^2}$*, times the expected (under the objective measure* P*) square root of the averaged quadratic variation of the log-price process plus a time-varying risk-premium, i.e.,*

$$\sigma_t^{BS} \approx \left(\sqrt{(1-\rho^2)}\right) \Im_t^P \left(\sqrt{U_{t,T}}\right) + RP_t. \tag{52}$$

Hence, if $\rho \neq 0$ as suggested by a substantial amount of recent evidence,[14] then the coefficient of the regression of realized on implied volatility should always be larger than one. In other words, if the reliance on a specific option pricing model is taken seriously, then the empirically verified correlation between stock returns and volatility should suggest more caution in interpreting the results that standard unbiasedness tests have furnished. Recent studies show an awareness of this problem (Chernov (2001)) but have a tendency to overlook its implications.[15] This is understandable in that accounting for the correlation between returns and volatility process might make the empirical investigation depend heavily on a tightly parametrized option pricing model, thus contradicting the largely model-free spirit of the exercise.

As in Section 2, we investigate the robustness of the long run one-to-one correspondence between implied and realized volatility series by estimating model (24) in the presence of the same two additional sets of data. The results (in the bottom two rows of Figure 4) reinforce our previous findings and are fully coherent with the simulations in the previous section.

The 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, use of least

[14] Following the intuition of Black (1975), i.e., "...a stock that drops sharply in price is likely to show a higher volatility in the future than a stock that rises sharply in price...", it is now widely accepted that there exists a negative correlation between stock returns and volatility, thus delivering stock return distributions that are negatively skewed (see Bakshi et al. (1997)). A non-zero correlation between the stock return process and volatility can explain the asymmetric smiles in the implied volatility curves (Renault (2001)).

[15] Chernov (2001) writes "the non-zero correlation will not affect the approximation in (2.3) ((56) in the present paper) by much."

squares would lead to rejection of the unbiasedness hypothesis. Furthermore, the downward trend in the point estimates is now much more evident and the confidence bands are substantially wider that in the case of our original data. We can single out two potential explanation for these effects. First, our original use of the VIX as the implied volatility measure (which, as discussed earlier, attempts to remove the effects of the American nature of options, dividends, nonsynchronous trading, and so on) 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 determining flatter slope estimates across different frequencies. Secondly, the VIX data set has a larger sample size leading to more precise estimation.

In general, our use of narrow band least squares 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 selection for spectral estimation with long-memory, but his work would have to be extended to the case of fractional cointegration. In effect, this choice 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. This trade-off was also apparent when inspecting our simulation experiments in the previous section.

One final observation is in order. 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 log transformation to be highest, followed by the standard deviation and, lastly, by the variance.[16]

7 Conclusion

This paper has argued that the persistence of measures of financial market volatility need to be taken into account when assessing the relation between realized and implied volatility. Coherently with recent work on the properties of realized volatility measures (see Andersen et al. (2001a, 2001b), among others), we use a simple modelling approach based on long-memory to capture persistence. In doing so, the features of the data suggest that the usual

[16] Corresponding figures can be provided by the authors upon request.

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 co-movements 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 methodology to estimate the standard regression between realized and implied volatility that is robust to the various issues that were raised in the extant literature as explanations for a slope coefficient less than one, i.e., measurement errors and presence of an unobservable time-varying risk premium correlated with the regression residuals, for example. In particular, our approach does not require to choose a particular parametrization for the risk premium and is thus robust to the likely mispecification that this choice would entail given that the existing asset pricing theory does not offer clear indications in this respect. More generally, our procedure is robust to various forms of probable non-orthogonality between the regressor (the implied volatility proxy) and the regression residuals.

Consistently with some recent studies (see Christensen and Prabhala (1998), Poteshman (2000) and Chernov (2001)), we do find evidence of near 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. In other words, we stress that little can be said about short-term unbiasedness, option market efficiency and/or validity of a certain option pricing model, which were the focus of much existing work on the subject.

This paper recognizes one possible instance where persistence of various volatility measures affect both statistical inference and economic interpretation. More generally, we think that care should be employed in empirical work any time a notion of volatility is believed to play a role in forecasting.

8 Proofs and Technical details

Proof of Proposition 1. Expanding the European option price C_t around $\sqrt{U_{t,T}}$, we can write

$$C_t = \Im_t^Q \begin{pmatrix} C^{BS}\left(S_t, \Im_t^Q\left(\sqrt{U_{t,T}}\right), r, K, T-t\right) \\ +\frac{\partial C}{\partial U_{t,T}}^{BS}\left(S_t, \Im_t^Q\left(\sqrt{U_{t,T}}\right), r, K, T-t\right)\left(\sqrt{U_{t,T}} - \Im_t^Q\left(\sqrt{U_{t,T}}\right)\right) \end{pmatrix} \tag{53}$$

$$+\Im_t^Q o\left(\left(\sqrt{U_{t,T}} - \Im_t^Q\left(\sqrt{U_{t,T}}\right)\right)\right) \tag{54}$$

$$= C^{BS}\left(S_t, \Im_t^Q\left(\sqrt{U_{t,T}}\right), r, K, T-t\right) + \Im_t^Q o\left(\left(\sqrt{U_{t,T}} - \Im_t^Q\left(\sqrt{U_{t,T}}\right)\right)\right) \tag{55}$$

and, in consequence,

$$\sigma_{imp} \approx \Im_t^Q\left(\sqrt{U_{t,T}}\right) = \Im_t^Q\left(\sqrt{\frac{1}{T-t}\int_t^T \sigma_s^2 ds}\right), \tag{56}$$

where the approximation is more accurate for at-the-money, short-term, options. Now notice that

$$\sigma_{imp}^2(t) \approx \Im_t^Q\left(\sqrt{U_{t,T}}\right) = \frac{\Im_t^P\left(\xi_T\sqrt{U_{t,T}}\right)}{\Im_t^P\left(\xi_T\right)} \tag{57}$$

$$= \frac{\Im_t^P\left(\xi_T\sqrt{U_{t,T}}\right)}{\xi_t} \tag{58}$$

$$= \Im_t^P\left(\frac{\xi_T}{\xi_t}\right)\Im_t^P\left(\sqrt{U_{t,T}}\right) + Cov_t^P\left(\frac{\xi_T}{\xi_t}, \sqrt{U_{t,T}}\right) \tag{59}$$

$$= \Im_t^P\left(\sqrt{U_{t,T}}\right) + Cov_t^P\left(\frac{\xi_T}{\xi_t}, \sqrt{U_{t,T}}\right) \tag{60}$$

$$= \Im_t^P\left(\sqrt{U_{t,T}}\right) + RP_t, \tag{61}$$

where ξ_t is the Radon-Nikodym derivative $\frac{dQ}{dP}$ of the risk-neutral measure Q with respect to the objective measure P. As usual, the positive-valued process ξ_t is defined as

$$\xi_t = \exp\left(-\int_0^t \boldsymbol{\lambda}_u^{'} d\mathbf{W}_u - \frac{1}{2}\int_0^t \boldsymbol{\lambda}_u^{\prime}\boldsymbol{\lambda}_u du\right)$$

with $\mathbf{W}_t = \left(W_t^1, W_t^2\right)^{'}$ and $\boldsymbol{\lambda}_t = \left(\lambda_t^1, \lambda_t^2\right)^{'}$, where $\boldsymbol{\lambda}_t$ is a vector of risk premia. Furthermore, ξ_t is a martingale under P provided the standard Novikov's condition, viz.,

$$\mathbf{E}\left(-\frac{1}{2}\int_0^T \boldsymbol{\lambda}_u^{\prime}\boldsymbol{\lambda}_u du\right) < \infty,$$

is satisfied. The martingale property is used in deriving equalities (58) and (59), whereas equality (57) follows from standard change of measure in the presence of conditional expectations. This proves the stated result.

Proof of Theorem 1. We employ the method of proof of Robinson (1994). Sufficiency can be shown easily. Using the Cauchy inequality, write

$$
\begin{aligned}
& \left|\widehat{\beta}_{n,m}-\beta\right| \\
= \; & \frac{\sum_{s=1}^{m}\varpi_{\varepsilon^*}(\lambda_s)\overline{\varpi_x(\lambda_s)}}{\sum_{s=1}^{m}\left|\varpi_x(\lambda_s)\right|^2} && (62)\\
= \; & \left|\frac{\sum_{s=1}^{m}\varpi_{u^R}(\lambda_s)\overline{\varpi_x(\lambda_s)}}{\sum_{s=1}^{m}\left|\varpi_x(\lambda_s)\right|^2}\right| + \left|\frac{\sum_{s=1}^{m}\varpi_{\varepsilon}(\lambda_s)\overline{\varpi_x(\lambda_s)}}{\sum_{s=1}^{m}\left|\varpi_x(\lambda_s)\right|^2}\right| \\
& + \left|\frac{\sum_{s=1}^{m}\varpi_{\delta}(\lambda_s)\overline{\varpi_x(\lambda_s)}}{\sum_{s=1}^{m}\left|\varpi_x(\lambda_s)\right|^2}\right| + \left|\frac{\beta\sum_{s=1}^{m}\varpi_{u^{BS}}(\lambda_s)\overline{\varpi_x(\lambda_s)}}{\sum_{s=1}^{m}\left|\varpi_x(\lambda_s)\right|^2}\right| && (63)\\
\leq \; & \left\{\frac{\sum_{s=1}^{m}\left|\varpi_{u^R}(\lambda_s)\right|^2}{\sum_{s=1}^{m}\left|\varpi_x(\lambda_s)\right|^2}\right\}^{1/2} + \left\{\frac{\sum_{s=1}^{m}\left|\varpi_{\varepsilon}(\lambda_s)\right|^2}{\sum_{s=1}^{m}\left|\varpi_x(\lambda_s)\right|^2}\right\}^{1/2} \\
& + \left\{\frac{\sum_{s=1}^{m}\left|\varpi_{\delta}(\lambda_s)\right|^2}{\sum_{s=1}^{m}\left|\varpi_x(\lambda_s)\right|^2}\right\}^{1/2} + |\beta|\left\{\frac{\sum_{s=1}^{m}\left|\varpi_{u^{BS}}(\lambda_s)\right|^2}{\sum_{s=1}^{m}\left|\varpi_x(\lambda_s)\right|^2}\right\}^{1/2} && (64)\\
\rightarrow \; & {}_p\left\{\frac{F_{\varepsilon\varepsilon}(1,m)}{F_{xx}(1,m)}\right\}^{1/2} + \left\{\frac{F_{u^R u^R}(1,m)}{\widehat{F}_{xx}(1,m)}\right\}^{1/2} \\
& + \left\{\frac{F_{\delta\delta}(1,m)}{F_{xx}(1,m)}\right\}^{1/2} + |\beta|\left\{\frac{F_{u^{BS}u^{BS}}(1,m)}{F_{xx}(1,m)}\right\}^{1/2} && (65)
\end{aligned}
$$

if $1/m + m/n \rightarrow 0$ as $n \rightarrow \infty$. But,

$$\frac{F_{\delta\delta}(1,m)}{F_{xx}(1,m)} = \frac{\int_0^{\lambda_m} f_{\delta\delta}(\theta)d\theta}{\int_0^{\lambda_m} f_{xx}(\theta)d\theta} \sim \frac{\frac{\lambda_m^{2(1/2-d_2)}}{2(1/2-d_2)}}{\frac{\lambda_m^{2(1/2-d_1)}}{2(1/2-d_1)}} \sim \lambda_m^{d_1-d_2} \rightarrow 0 \tag{66}$$

since $\lambda_m \rightarrow 0$ when $n \rightarrow \infty$, given $d_1 > d_2$. Then, the bound becomes

$$c_1\lambda_m^{d_1-d_2} + c_2\lambda_m^{d_1} \rightarrow 0 \tag{67}$$

which, in turn, implies

$$\left|\widehat{\beta}_{n,m}-\beta\right| \xrightarrow{p} 0. \tag{68}$$

Now, write the estimation error decomposition for $\widehat{\alpha}_{n,m}$ as

$$\widehat{\alpha}_{n,m} - \alpha = -\left(\widehat{\beta}_{n,m}-\beta\right)\sqrt{\frac{2\pi}{n}}\varpi_y(0) + \sqrt{\frac{2\pi}{n}}\varpi_{\varepsilon^*}(0). \tag{69}$$

Note that

$$\sqrt{\frac{2\pi}{n}}\varpi_{\varepsilon^*}(0) \rightarrow_p 0 \tag{70}$$

by the Ergodic Theorem. Combining (68) and (70), it follows that $\widehat{\alpha}_{n,m} \xrightarrow{p} \alpha$. This proves the stated result.

Proof of Proposition 2. By Ito's lemma, we can write the dynamic evolution of the log-price process as

$$d\log S_t = \left(r_t - \frac{\sigma_t^2}{2}\right)dt + \rho\sigma_t dW_t^{2*} + \sqrt{1-\rho^2}\sigma_t dW_t^{1*}. \tag{71}$$

As a consequence, conditionally on $\Im_t$, the volatility path $\{\sigma_t\}_{t\in[0,T]}$ and the path of $\left\{dW_t^{2*}\right\}_{t\in[0,T]}$, the distribution of $\log\left(\frac{S_T}{S_t}\right)$ is normal with mean

$$\int_t^T r_t dt - \left(1-\rho^2\right)\frac{\overline{U}_{t,T}}{2} - \rho^2\frac{\overline{U}_{t,T}}{2} + \rho\int_t^T \sigma_t dW_t^{2*}, \tag{72}$$

where

$$\overline{U}_{t,T} = \int_t^T \sigma_t^2 dt \tag{73}$$

and variance

$$\left(1-\rho^2\right)\overline{U}_{t,T}. \tag{74}$$

This fact implies that the price C_t of a European call option with strike K and expiration T can be written as

$$\begin{aligned} C_t &= \Im_t^Q\left(\exp\left(-\int_t^T r(s)ds\right)(S_T-K)^+\right) && (75)\\ &= \Im_t^Q\left(\Im_t^Q\left(\exp\left(-\int_t^T r_s ds\right)(S_T-K)^+ | W^{2*}, \{\sigma_t\}_{t\in[0,T]}\right)\right) && (76)\\ &= \Im_t^Q\left(C^{BS}\left(S_t V_{t,T}, \sqrt{U_{t,T}}, K, T-t\right)\right), && (77)\end{aligned}$$

where

$$V_{t,T} = \exp\left(-\rho^2\frac{\overline{U}_{t,T}}{2} + \rho\int_t^T \sigma_t dW_t^{2*}\right) \tag{78}$$

implying

$$\Im_t^Q[V_{t,T}] = 1 \tag{79}$$

and

$$U_{t,T} = (1-\rho^2)\frac{\overline{U}_{t,T}}{T-t}. \tag{80}$$

Using a similar approximation as that in the proof of Proposition 2, but taking into account the fact that such approximation can be substantially worse that earlier due to the fact that $V_{t,T}$ equals one only in expectation, we can now write

$$\sigma_{imp} \approx \Im_t^Q\left(\sqrt{U_{t,T}}\right) = \Im_t^Q\left(\sqrt{\frac{(1-\rho^2)}{T-t}\int_t^T \sigma^2(s)ds}\right) \tag{81}$$

and, consequently,

$$\Im_t^Q\left(\sqrt{U_{t,T}}\right) \approx \frac{1}{(1-\rho^2)}\sigma_{imp}. \tag{82}$$

This proves the stated result.

References

[1] Akonom, J. and C. Gourieroux, "A Functional Central Limit Theorem for Fractional Processes," Technical Report #8801, CEPREMAP, Paris, 1987.

[2] Alizadeh, S., M. W. Brandt, and F. X. Diebold, "Range-Based Estimation of Stochastic Volatility Models," *Journal of Finance*, 57, June 2002.

[3] Andersen, T. G. and Tim Bollerslev, "Answering the Critics: Yes, ARCH Models Do Provide Good Volatility Forecasts," *International Economic Review*, 39, 1998, 885-905.

[4] Andersen, T. G., T. Bollerslev, F. X. Diebold, and P. Labys, "The Distribution of Realized Exchange Rate Volatility," *Journal of American Statistical Association*, 2001a, 96, 42-55.

[5] 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.

[6] Andrews, D. W. K. and P. Guggenberger, "A Bias-Reduced Log-Periodogram Regression Estimator for the Long Memory Parameter," Mimeo, Yale University, 2000.

[7] Bakshi, G., C. Cao and Z. Chen, "Empirical Performance of Alternative Option Pricing Models," *Journal of Finance,* 1997, 2003-2049.

[8] Baillie, R., "Long Memory Processes and Fractional Integration in Econometrics," *Journal of Econometrics,* 73, 1996, 5-59.

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

[10] Barnhart, S.W. and A. Szakmary, "Testing the Unbiasedness Forward Rate Hypothesis: Evidence on Unit Roots, Co-integration and Stochastic Coefficients," *Journal of Financial and Quantitative Analysis*, 26, 1991, 245-267.

[11] Benzoni, L., "Pricing Options under Stochastic Volatility: An Econometric Analysis," Working Paper, Northwestern University, 1998.

[12] Beran, J., *Statistics for Long Memory Processes*, Chapman and Hall, 1994.

[13] 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.

[14] Black, F., "Fact and Fantasy in the Use of Options," *Financial Analysts Journal,* 1975, 36-72.

[15] Bollerslev, T. and H. O. Mikkelsen, "Long-term Equity Anticipation Securities and Stock Market Volatility Dynamics," *Journal Of Econometrics*, 92, 1999, 75-99.

[16] Bollerslev, T. and H.O. Mikkelsen, "Modeling and Pricing Long-Memory in Stock Market Volatility," *Journal of Econometrics*, 73, 1996.

[17] Brunetti, C. and C.L. Gilbert, "Bivariate FIGARCH and Fractional Cointegration," *Journal of Empirical Finance,* 7, 2000, 509-530.

[18] Canina, L. and S. Figlewski, "The Informational Content of Implied Volatility," *The Review of Financial Studies*, 6, 1993, 659-681.

[19] Chernov, M., "Implied Volatilities as Forecasts of Future Volatility, the Market Risk Premia, and Returns Variability," Mimeo, Columbia University, 2000.

[20] Chernov, M., A. R. Gallant, E. Ghysels, and G. Tauchen, "Alternative Models for Stock Price Dynamics," Working paper, 2001.

[21] Christensen, B.J. and N.R. Prabhala, "The Relation Between Implied and Realized Volatility," *Journal of Financial Economics,* 50, 1998, 125-150.

[22] Comte, F., "Simulation and Estimation of Long Memory Continuous-Time Models," *Journal of Time Series Analysis*, 17, 1996, 19-36.

[23] Comte, F. and E. Renault, "Long-Memory Continuous-Time Models," *Journal of Econometrics,* 73, 1996, 101-149.

[24] Comte, F. and E. Renault, "Long-Memory in Continuous-Time Stochastic Volatility Models," *Mathematical Finance,* 8, 1998, 291-323.

[25] 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

[26] Day, T. and C. Lewis, "Stock Market Volatility and the Information Content of Stock Index Options," *Journal of Econometrics,* 52, 1992, 267-287.

[27] Day, T. and C. Lewis, "Forecasting Futures Market Volatility," *Journal of Derivatives,* 1, 1993, 33-50.

[28] Davydov, Y., "The Invariance Principle for Stationary Processes," *Theory of Probability and Its Applications,* 15, 1970, 487-489

[29] 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.

[30] Dittmann, I. and C. W. J. Granger, "Properties of Nonlinear Transformations of Fractionally Integrated Processes," Discussion Paper 2000-07, Department of Economics, University of California, San Diego, 2000.

[31] Ebens, H., "Realized Stock Volatility," Working Paper, The Johns Hopkins University, 1999.

[32] Engle, R. F. and G. J. Lee, "A Permanent and Transitory Component Model of Stock Return Volatility", in Engle, R. F. and H. White, eds. *Cointegration, Causality, and Forecasting: A Festschriftin Honor of Clive W.J. Granger*, Oxford University Press, 1999, 475-497.

[33] Eraker, B., M. Johannes, and N. Polson, "The Impact of Jumps in Returns and Volatility," *Journal of Finance,* forthcoming, 2001.

[34] Fleming, J., "The Quality of Market Volatility Forecasts Implied by S&P 100 Index Option Prices," *Journal of Empirical Finance,* 5, 1998, 317-354.

[35] French, K.R., G.W. Schwert, and R.F. Stambaugh, "Expected Stock Returns and Volatility," *Journal of Financial Economics*, 19, 3-29.

[36] Gallant, A. R. and G. Tauchen, "EMM: A Program for Efficient Method of Moments Estimation, version 1.6,", Mimeo, Duke University, 2002.

[37] Gallant, A. R. and G. Tauchen, "Which Moments to Match?", *Econometric Theory*, 12, 1996, 657-681.

[38] 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.

[39] 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.

[40] Hidalgo, J., "An Alternative Bootstrap to Moving Blocks for Time Series Regression Models," Mimeo, LSE, March 2001.

[41] Hull, J. and A. White, "The Pricing of Options on Assets with Stochastic Volatilities," *Journal of Finance,* 42, 1987, 281-301.

[42] Hurvich, C.M., R. Deo and J. Brodsky, "The Mean Squared Error of Geweke and Porter-Hudak's Estimator of The Memory Parameter of a Long Memory Time Series," *Journal of Time Series Analysis,* 19, 1998, 19-46.

[43] Jackwerth, J. C. and M. Rubinstein, "Recovering Probability Distributions from Option Prices," *Journal of Finance,* 51, 2000, 1611-1631.

[44] Jorion, P., "Predicting Volatility in the Foreign Exchange Market," *The Journal of Finance*, 50, 1995, 507-528.

[45] Kim, C. S. and P. C. B. Phillips, "Log Periodogram Regression: The Nonstationary Case," Mimeo, Yale University, 1999.

[46] Kim, C. S. and P. C. B. Phillips, "Fully Modified Estimation of Fractional Cointegration Models," Mimeo, Yale University, 1999b.

[47] Lamoureux, D. and W. Lastrapes, "Forecasting Stock-Return Variance: Toward an Understanding of Stochastic Implied Volatilities," *Review of Financial Studies,* 6, 1993, 293-326.

[48] Li, K., "Yes, Historical Volatility Does Contain Incremental Information Beyond Option Implied Volatility," Working Paper, Stern School of Business, New York University, 2000.

[49] Marinucci, D. and P.M. Robinson, "Alternative Forms of Fractional Brownian Motion," *Journal of Statistical Planning and Inference*, 80, 1999, 111-122.

[50] Marinucci, D. and P.M. Robinson, "Weak Convergence of Multivariate Fractional Processes," *Stochastic Processes and their Applications*, 86, 2000, 103-120.

[51] Pagan, A. R. and G. W. Schwert, "Alternative Models for Conditional Stock Volatility," *Journal of Econometrics*, 45, 1990, 267-290.

[52] Pan, J., "The Jump-Risk Premia Implicit in Options: Evidence from an Integrated Time-Series Analysis," *Journal of Financial Economics*, 63, 2002, 3-50.

[53] Park, J.Y. and P. C.B. Phillips, "Statistical Inference in Regressions with Integrated Processes: Part 1," *Econometric Theory,* 4, 1988, 468-497.

[54] Pham, H. and N. Touzi, "Intertemporal Equilibrium Risk Premia in a Stochastic Volatility Model," *Mathematical Finance,* 6, 1996, 215-236.

[55] Politis, D.N., and J.P. Romano, "Large Sample Confidence regions Based on Subsamples under Minimal Conditions," *Annals of Statistics*, 22, 1994, 2031-2050.

[56] Politis, D.N., J.P. Romano, and M. Wolf, *Subsampling*, Springer-Verlag. New York, 1999.

[57] Poteshman, A. M., "Forecasting future Volatility from Option Prices," Mimeo, University of Illinois at Urbana-Champaign, 2000.

[58] Protter, P., *Stochastic Integration and Differential Equations,* Springer-Verlag, New York, 1990.

[59] 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.

[60] Renault, E., "Econometric Methods for Option Pricing: Lecture Notes," Workshop of Financial Mathematics and Econometrics, Montreal, June 26-30, 2001.

[61] Robinson, P. M., "Semiparametric Analysis of Long-Memory Time Series," *Annals of Statistics*, 22, 1994, 515-539.

[62] Robinson, P. M., "Rates of Convergence and Optimal Spectral Bandwidth for Long Range Dependence," *Probability Theory and Related Fields*, 99, 1994b, 443-473.

[63] Robinson, P. M., "Log-Periodogram Regression of Time Series with Long Range Dependence," *Annals of Statistics*, 23, 1995, 1048-1072.

[64] Robinson, P. M., "Gaussian Semiparametric Estimation of Long Range Dependence," *Annals of Statistics*, 23, 1995b, 1630-1661.

[65] Robinson, P. M. and D. Marinucci, "Narrow-band Analysis of Nonstationary Processes", Mimeo, LSE, 2001.

[66] Rogers, L.C.G., "Arbitrage with Fractional Brownian Motion," *Mathematical Finance,* 7, 1997, 95-105.

[67] Samorodnitsky, G. and M.S. Taqqu, Stable Non-Gaussian Random Processes, Chapman and Hall, New York,1994.

[68] Schwert, G.W., "Stock Volatility in the New Millenium: How Wacky is Nasdaq?" *Journal of Monetary Economics*, 49, 2002, 1-24.

[69] Schwert, G.W., "Why Does Stock Market Volatility Change Over Time?" *Journal of Finance*, 44, 1989, 1115-1153.

[70] Silveira, G., Contributions to Strong Approximations in Time Series with Applications in Nonparametric Statistics and Functional Central Limit Theorems, Ph.D. thesis, University of London, 1991.

[71] Taqqu, M.S., "Convergence of Integrated Processes of Arbitrary Hermite Rank," *Z. Wahrscheinlichkeitstheorie und verwandte Gebiete,* 50, 1979, 53-83.

[72] Taylor, S., *Modeling Financial Time Series*, Wiley, Chichester, 1986.

[73] Velasco, C., "Non-stationary log-periodogram Regression," *Journal of Econometrics,* 91, 1999, 325-371.

Table 1.

Descriptive statistics for the realized and implied volatility series employed in this study. We use monthly data on the S&P 100 index realized standard deviation - σ^R - and the squared root of the VIX (the CBOE Market Volatility Index) - σ^{BS} - from January 1988 to May 2000 (152 non-overlapping observations). We compute the realized standard deviation over the remaining 30 days (one month) of the option as

$$\sigma^R_{t,30} = \sqrt{\frac{1}{30}\sum_{j=1}^{30} r^2_{t+j}},$$

where $r_t = \log(S_t/S_{t-1})$ and S_t is the daily S&P 100 index at date t. We report mean, standard deviation, skewness and kurtosis for both volatility measures as well as for two linear combinations of the same measures, namely $\sigma^R - \sigma^{BS}$ and $\sigma^R - \widehat{\alpha} - \widehat{\beta}\sigma^{BS}$, where $\widehat{\alpha}$ and $\widehat{\beta}$ are least squares parameter estimates.

The skewness and kurtosis coefficients are computed after studentizing the relevant quantity, say θ, as $t_\theta = \frac{(\theta_t - \overline{\theta})}{\sigma(\theta)}$, where $\sigma(\theta)$ is the standard deviation of θ. Hence, they are equal to $\frac{1}{n-1}\sum t^3_\theta$ and $\frac{1}{n-1}\sum t^4_\theta$, respectively.

	Mean	Std. dev.	Skewness	Kurtosis
Implied volatility (σ^{BS})	.1644	.0526	1.05	5.15
Realized volatility (σ^R)	.1434	.0621	1.46	6.10
Difference ($\sigma^R - \sigma^{BS}$)	-.0210	.0410	1.34	6.71
LS residuals ($\sigma^R - \widehat{\alpha} - \widehat{\beta}\sigma^{BS}$)	.0000	.0406	1.59	7.22

Table 2.

Summary of the simulation results as described in section 5. The log price process is generated according to the discretized Comte and Renault (1998) model, i.e.,

$$\ln S_{t+\Delta} = \ln S_t + \left(r - \frac{\sigma_t^2}{2}\right)\Delta + \sqrt{\Delta}\sigma_{t+\Delta}\varepsilon_{t+\Delta}^1 \tag{83}$$

$$\ln \sigma_t = \theta + \sum_{j=1}^{[t/\Delta]} \frac{(t - j\Delta)^d}{\Gamma(1+d)} \left[x_{j\Delta}^{(d)} - x_{(j-1)\Delta}^{(d)}\right] \tag{84}$$

$$x_{t+\Delta}^{(d)} = e^{-k\Delta} x_t^{(d)} + \gamma \left(\frac{e^{-2k\Delta} - 1}{-2k}\right)^{\frac{1}{2}} \varepsilon_{t+\Delta}^2 \tag{85}$$

$$\begin{pmatrix} \varepsilon_{t+\Delta}^1 \\ \varepsilon_{t+\Delta}^2 \end{pmatrix} \sim iidN(0, I_2) \tag{86}$$

with $r = 0.00962$, $\theta = -4.673$, $k = 1.561$, $\gamma = 0.388$ and $d = 0.506$, where the parameters are estimated by EMM (Gallant and Tauchen (1996)) using daily data on the S&P 100 index. We set $x_0^{(d)} = 0$ and normalize the log price process to be 0 at time 0. We take Δ to be half a day. Each replication involves computing 152 call option prices (numerically and on the basis of 20,000 simulations) that we invert to obtain the BS implied volatilities. Specifically, for each replication we generate a series of $2 \times 22 \times 152 = 6,688$ realizations of the log price process plus 1,000 points to be discarded. Realized volatility is computed by summing the squares of the daily returns over the life of the option, i.e., 22 squared returns corresponding to every other simulated value of the log price process. The path of the stock price used to compute the realized volatility series is chosen at random among the 20,000 simulations. Finally, we contaminate the implied volatility series by adding an error term with mean zero and variance equal to the sample variance of the implied volatility series over the corresponding replication. We repeat the procedure 1,000 times.

For each replication we run a linear regression between realized and implied volatility. We estimate the model by ordinary least squares (second row) and narrow band least squares. The number of Fourier frequencies used to compute the narrow band estimates is equal to the integer part of $n^{0.5}$, $n^{0.6}$, $n^{0.7}$ and $n^{0.8}$. The means of the resulting estimates (with standard errors underneath) are as follows:

	α	β	R^2 (%)
OLS	0.062	0.354	5.34
	(0.202)	(2.138)	
NBLS $\left(m = n^{0.5}\right)$	-0.027	0.903	
	(0.559)	(5.723)	
NBLS $\left(m = n^{0.6}\right)$	-0.010	0.797	
	(0.451)	(4.604)	
NBLS $\left(m = n^{0.7}\right)$	0.015	0.606	
	(0.350)	(3.563)	
NBLS $\left(m = n^{0.8}\right)$	0.034	0.464	
	(0.260)	(2.643)	

Table 3.

Summary of the simulation results as described in section 5. The log price process is generated according to the discretized Comte and Renault (1998) model, i.e.,

$$\ln S_{t+\Delta} = \ln S_t + \left(r - \frac{\sigma_t^2}{2}\right)\Delta + \sqrt{\Delta}\sigma_{t+\Delta}\varepsilon_{t+\Delta}^1 \tag{87}$$

$$\ln \sigma_t = \theta + \sum_{j=1}^{[t/\Delta]} \frac{(t - j\Delta)^d}{\Gamma(1+d)} \left[x_{j\Delta}^{(d)} - x_{(j-1)\Delta}^{(d)}\right] \tag{88}$$

$$x_{t+\Delta}^{(d)} = e^{-k\Delta} x_t^{(d)} + \gamma \left(\frac{e^{-2k\Delta} - 1}{-2k}\right)^{\frac{1}{2}} \varepsilon_{t+\Delta}^2 \tag{89}$$

$$\begin{pmatrix} \varepsilon_{t+\Delta}^1 \\ \varepsilon_{t+\Delta}^2 \end{pmatrix} \sim iidN(0, I_2) \tag{90}$$

with $r = 0.00962$, $\theta = -4.673$, $k = 1.561$, $\gamma = 0.388$ and $d = 0.506$, where the parameters are estimated by EMM (Gallant and Tauchen (1996)) using daily data on the S&P 100 index. We set $x_0^{(d)} = 0$ and normalize the log price process to be 0 at time 0. We take Δ to be half a day. Each replication involves computing 152 call option prices (numerically and on the basis of 20,000 simulations) that we invert to obtain the BS implied volatilities. Specifically, for each replication we generate a series of $2 \times 22 \times 152 = 6,688$ realizations of the log price process plus 1,000 points to be discarded. Realized volatility is computed by summing the squares of the daily returns over the life of the option, i.e., 22 squared returns corresponding to every other simulated value of the log price process. The path of the stock price used to compute the realized volatility series is chosen at random among the 20,000 simulations. Finally, we contaminate the implied volatility series by adding an error term with mean zero and variance equal to the sample variance of the implied volatility series over the corresponding replication. We repeat the procedure 1,000 times.

For each replication we estimate the long memory parameter d for both implied and realized volatility using the Geweke-Porter-Hudak (GPH) and the Andrews-Guggenberger (AG) estimator with a number of Fourier frequencies included in the appropriate regression equal to the integer part of $n^{0.5}$, $n^{0.6}$, $n^{0.7}$ and $n^{0.8}$. The means of the resulting estimates (with standard errors underneath) are as follows:

	$\hat{d}$ (implied)		$\hat{d}$ (realized)	
	GPH	AG	GPH	AG
$m = n^{0.5}$	0.452	0.437	0.567	0.704
	(0.308)	(0.529)	(0.256)	(0.485)
$m = n^{0.6}$	0.400	0.478	0.449	0.640
	(0.230)	(0.369)	(0.184)	(0.311)
$m = n^{0.7}$	0.320	0.452	0.339	0.524
	(0.173)	(0.272)	(0.140)	(0.216)
$m = n^{0.8}$	0.252	0.364	0.257	0.397
	(0.130)	(0.203)	(0.104)	(0.163)

Fig. 1. d estimates for realized and implied - S&P 100 and VIX

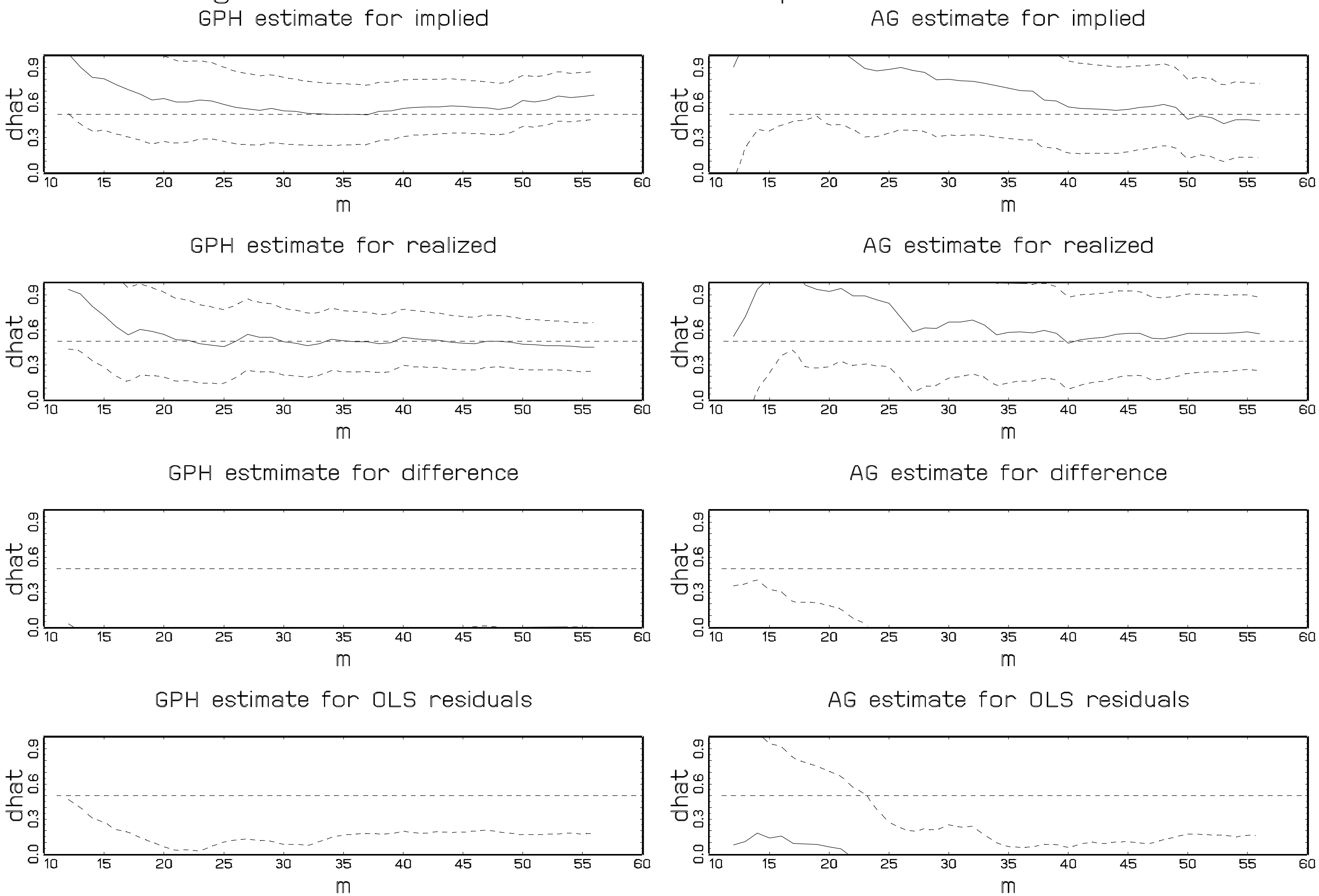

Fig. 2. d estimates for realized and implied – S&P 100

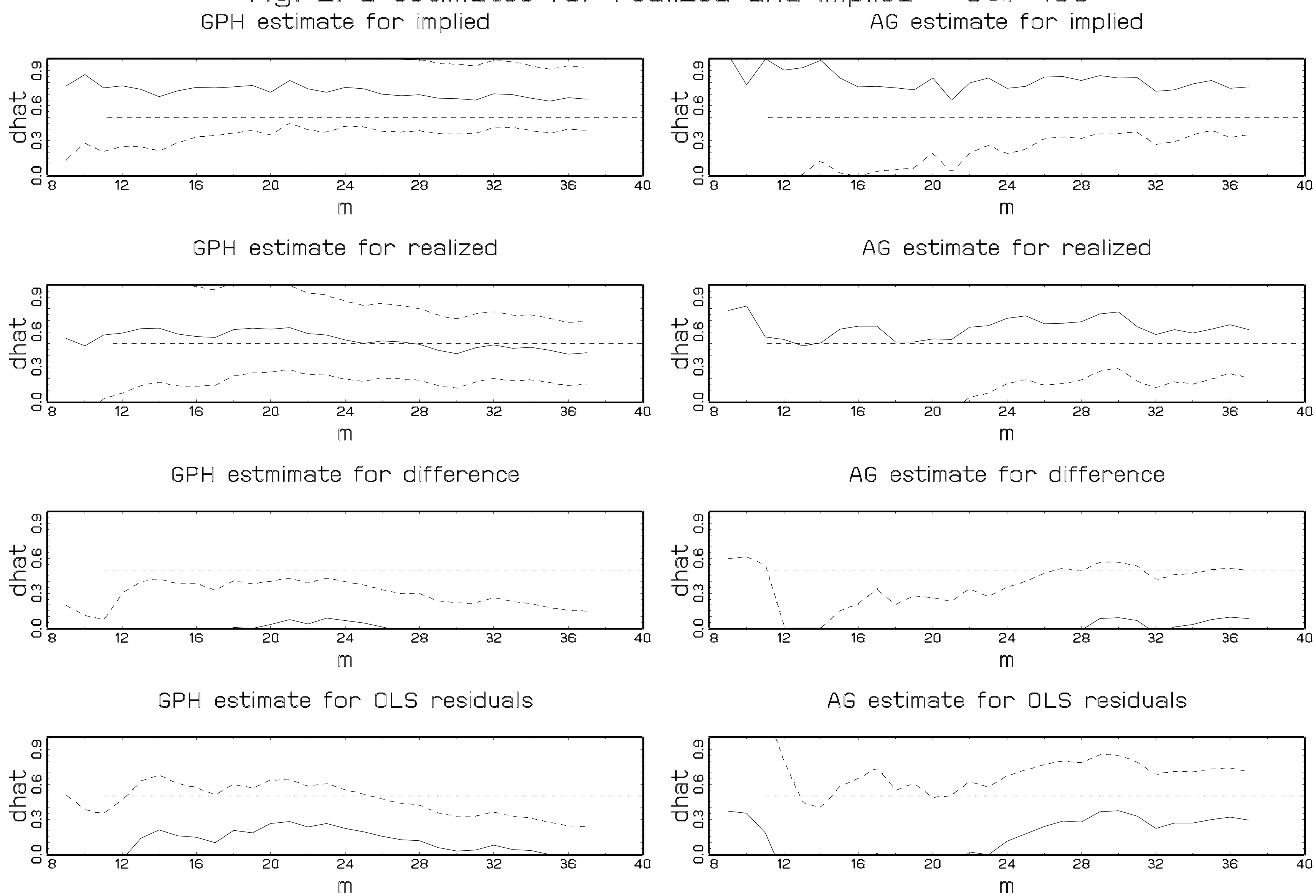

Fig. 3. d estimates for realized and implied - DM data

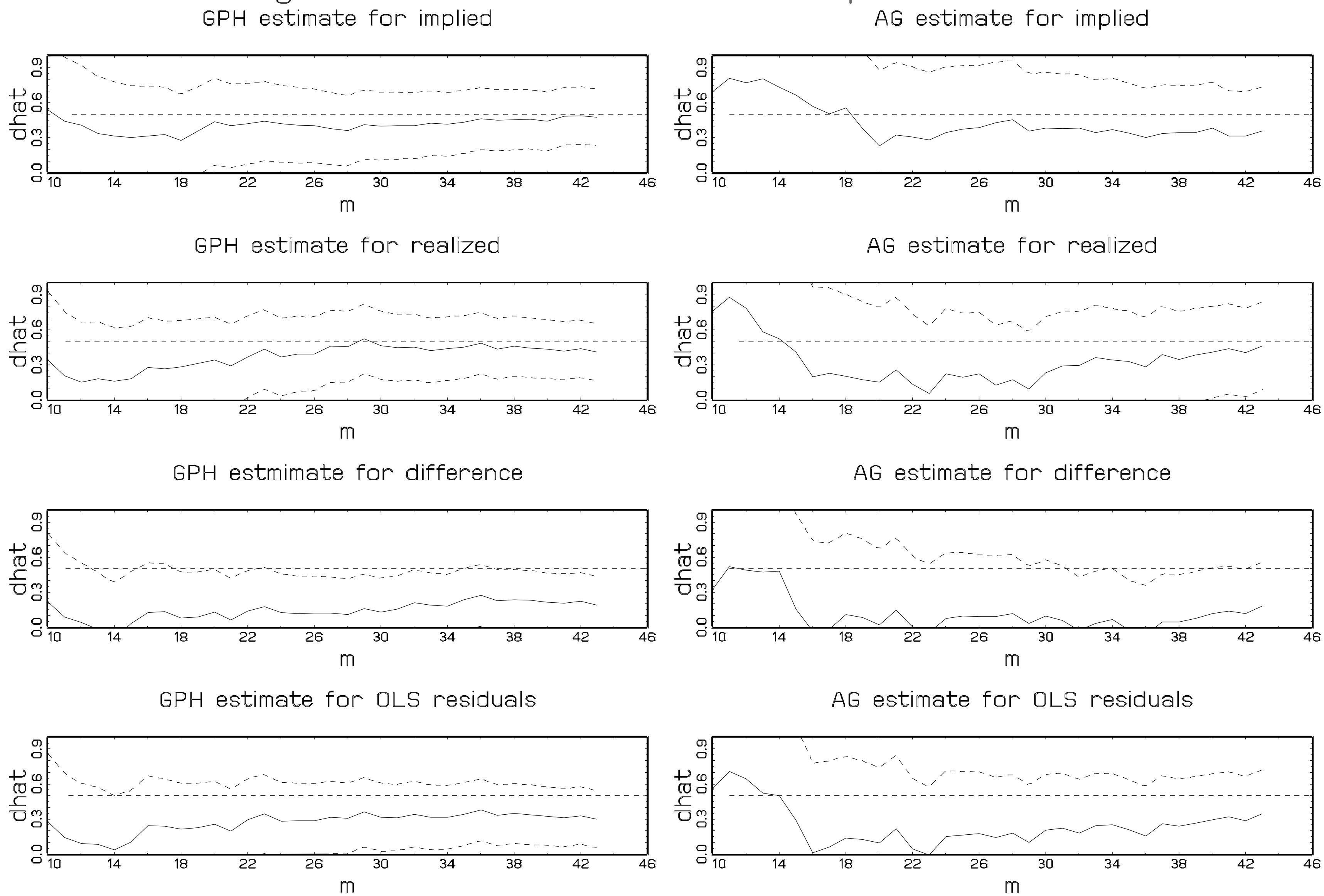

Fig. 4. Narrow band estimates of constant and slope

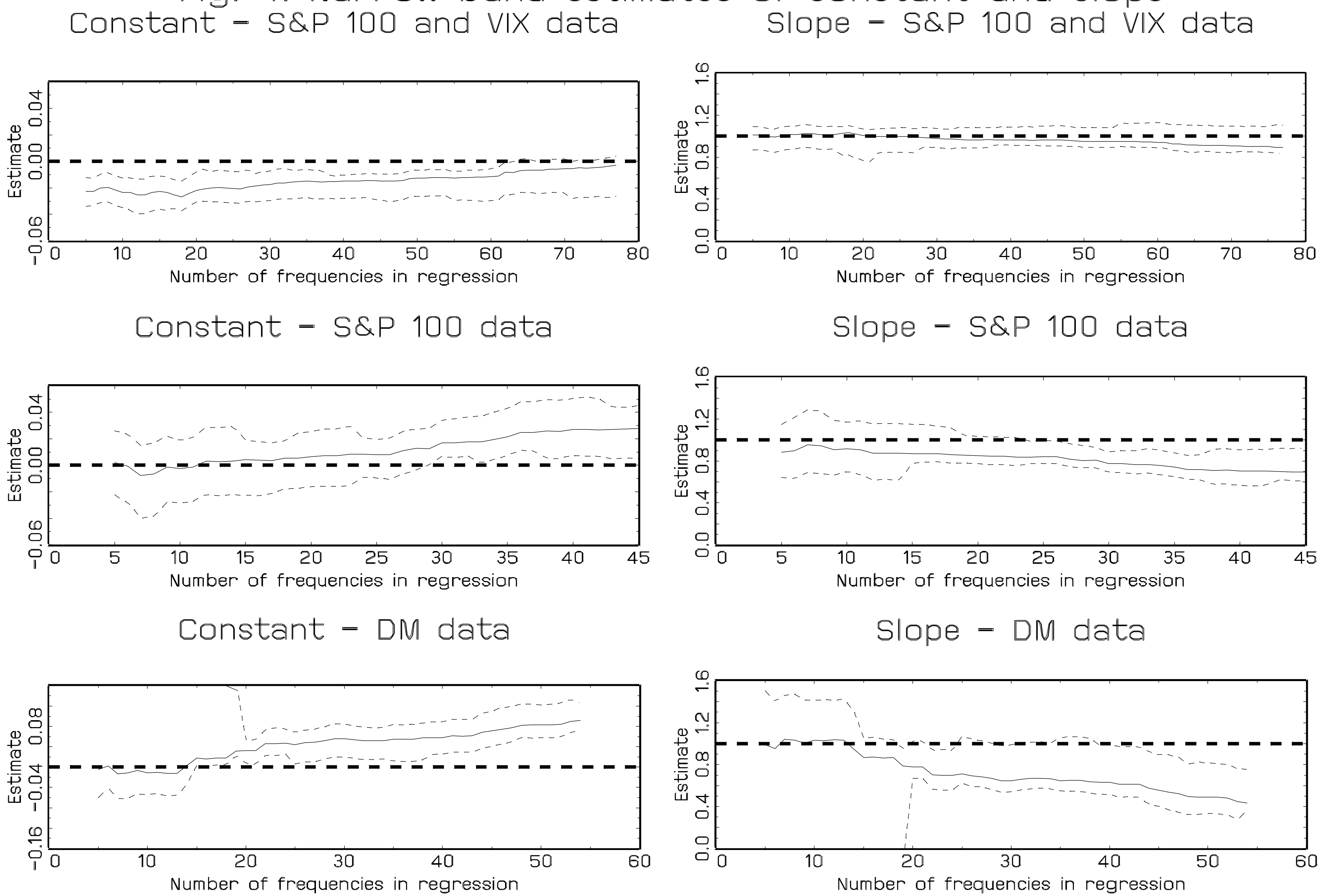

Fig. 4. Narrow band estimates of constant and slope