REALIZED VOLATILITY FORECASTING in the PRESENCE of TIME-VARYING NOISE*

(Preliminary - Comments Welcome)

Federico M. Bandi, Jeffrey R. Russell, Chen Yang
Graduate School of Business, University of Chicago

First draft: November 2006
This draft: January 2007

Abstract

Observed high-frequency financial prices can be considered as having two components, a true price and a market microstructure noise perturbation. It is an empirical fact that the second moment of market microstructure noise is time-varying. We study the optimal design of nonparametric variance estimators in linear forecasting models with time-varying market microstructure noise. Specifically, we discuss optimal frequency selection in the case of the classical realized variance estimator and optimal bandwidth selection in the case of kernel-type integrated variance estimators. In this setting, we show that the sampling frequencies are generally considerably lower (the bandwidths are generally considerably larger) than those that would be optimally chosen in linear forecasting models when time-variation in the second moment of the noise is unaccounted for. Conditional and unconditional frequency/bandwidth choices are discussed.

Keywords: Realized Variance, Kernel-based Estimators, Time-Varying Market Microstructure Noise, Volatility Forecasting.

*We thank Torben Andersen and Nour Meddahi for discussions. We are grateful to the William S. Fishman Faculty Research Fund at the Graduate School of Business of the University of Chicago (Bandi), NYU Stern (Russell), and the Graduate School of Business of the University of Chicago (Bandi, Russell, and Yang) for financial support. Correspondence can be sent to: federico.bandi@chicagogsb.edu, jeffrey.russell@chicagogsb.edu, cyang1@chicagogsb.edu.

1 Introduction

Observed high-frequency financial prices can be considered as having two components, a true price and a market microstructure noise perturbation. It is an empirical fact that the second moment of market microstructure noise is time-varying (for evidence, Bandi and Russell, 2006b, Hansen and Lunde, 2006, and Oomen, 2006). This time variation (as illustrated in Figure 1) induces time variation in the bias of the realized variance estimator constructed using high-frequency data. Naturally, a time-varying bias in realized variance has implications for variance forecasting. This paper shows that the time-varying nature of market microstructure noise is a fundamental aspect of the variance forecasting problem. This is true both from a theoretical and an empirical perspective.

Optimal frequency selection (in the case of realized variance) and bandwidth selection (in the case of alternative kernel-type variance estimators) have been very active areas of recent research (Bandi and Russell, 2006c, Barndorff-Nielsen and Shephard, 2006, and McAleer and Medeiro, 2006, are recent surveys on the subject). Joint consideration of the frequency/bandwidth selection and the forecasting problem is pursued by Andersen et al. (2006), ABM hereafter, and Ghysels and Sinko (2006), GS henceforth. Specifically, in the context of theoretical linear regression models with time-invariant noise, they show that the optimal frequency/bandwidth selection problem (for the purpose of R^2 maximization) reduces to minimization of the unconditional variance of the estimator (regressor).

We derive novel optimal frequency and bandwidth selection methods for the joint frequency/bandwidth selection and forecasting problem in the presence of time-varying microstructure noise variance. In this context, we find considerably lower optimal frequencies and, similarly, considerably larger bandwidths than those obtained when the time variation in the noise variance is unaccounted for. Interestingly, the new frequency/bandwidth choices are in general closer to those that would be derived from optimization of the finite sample unconditional MSE of the regressor (as in Bandi and Russell, 2006a, 2006b) than to those that would be obtained from optimization of the unconditional variance of the regressor (under the assumption of a constant bias term), as needed in linear forecasting models with time-invariant noise (ABM, 2006, GS, 2006). Specifically, we find that taking bias into account *suboptimally* through an unconditional MSE-based optimization (under the assumption of a constant bias) can be an empirically reasonable strategy.

Finally, we discuss cases in which choosing the sampling frequency/bandwidth conditionally (for each entry/day in the regressor vector) is a superior strategy from a forecasting standpoint than making the same choice unconditionally, even when the latter is the exact, unconditional, R^2- optimal choice. An application to SPIDERS mid-quotes validates our theoretical predictions.

Several promising, recent contributions have studied variance forecasting using microstructure noise-contaminated high-frequency variance estimates. These contributions have evaluated the forecasting potential of the classical realized variance estimator (Andersen et al., 2003, and Barndorff-Nielsen et al., 2002) as well as that of classes of more robust (to noise) kernel-based estimators (e.g., Zhou, 1996, Zhang et al., 2005, and Barndorff-Nielsen et al., 2006). Given time-series of alternative variance estimates, the forecasts have generally been obtained using ARFIMA models (Bandi and Russell, 2005, 2006a, Bandi et al., 2006), Mincer-Zarnowitz-style linear regressions (Andersen et al., 2006), or MIDAS-type regressions (Ghysels and Synko, 2006a, 2006b). Either statistical metrics, such as forecast mean-squared errors and coefficients of determination (Aït-Sahalia and Mancini, 2006, Andersen et al., 2006, Corradi et al., 2006, and Ghysels and Sinko, 2006a, 2006b) or economic metrics, such as the utility obtained by investors or the profits obtained by option traders on the basis of alternative variance forecasts (Bandi and Russell, 2005, 2006a, Bandi et al., 2006), have been used for the purpose of evaluating the goodness of the forecasts.

Time-varying noise variance and its impact on variance forecasting have been considered by Bandi and Russell (2005, 2006a) and Bandi et al. (2006) using *conditional* frequency/bandwidth selection rules. The joint frequency/bandwidth selection and forecasting problem has been studied by ABM (2006) and GS (2006) using *unconditional* optimal rules in the presence of a time-invariant second moment of the noise. In this paper, we study conditional and unconditional optimality methods for the joint problem in the presence of time-varying noise.

2 The Price Formation Mechanism

Consider a trading day t. Assume availability of $M+1$ observed logarithmic asset prices over $[t, t+1]$ and write

$$p_{t+j\delta} \;=\; p^*_{t+j\delta} + \eta_{t+j\delta} \qquad j = 0, ..., M$$

or, in terms of continuously-compounded returns,

$$\underbrace{p_{t+j\delta} - p_{t+(j-1)\delta}}_{r_{t+j\delta}} \;=\; \underbrace{p^*_{t+j\delta} - p^*_{t+(j-1)\delta}}_{r^*_{t+j\delta}} + \underbrace{\eta_{t+j\delta} - \eta_{t+(j-1)\delta}}_{\varepsilon_{t+j\delta}}, \qquad j = 1, ..., M,$$

where p^* denotes the *unobservable* equilibrium price, η denotes *unobservable* market microstructure noise, and $\delta = \frac{1}{M}$ represents the time distance between adjacent price observations.

We assume the equilibrium price process evolves in time as a stochastic volatility local martingale, i.e.,

$$p^*_t = \int_0^t \sigma_s dW_s,$$

where σ_t is a càdlàg stochastic volatility process and W_t is a standard Brownian motion. We also assume the noise contaminations in the price process η are independent of p^*_t and IID over each day with variance σ^2_{tu} (independent of the equilibrium price integrated variance $V_t = \int_t^{t+1} \sigma^2_s ds$) and fourth moment $K_u \sigma^2_{tu}$.

Importantly, the variance of the market microstructure noise carries a subscript t to indicate that it can change from day to day.

3 Forecasting with Realized Variance

We are interested in predicting V_{t+1} given past values of the classical realized variance estimator, namely $\widehat{V}_t = \sum_{j=1}^{M} r^2_{t+j\delta}$ (Andersen et al., 2003, and Barndorff-Nielsen et al., 2002). Assume $V_{t+1} = \alpha + \beta V_t + \varepsilon_{t+1}$, where ε_{t+1} is such that $\mathbf{E}(\varepsilon_{t+1}|\mathfrak{F}_t) = 0$, and consider the forecasting regression

$$V_{t+1} = \widehat{\alpha} + \widehat{\beta}\widehat{V}_t + \widehat{\varepsilon}_{t+1}. \tag{1}$$

The following theorem presents the optimal rule to choose the R^2-maximizing number of observations M.

Theorem 1.

Consider the regression in Eq. (1). *Then,*

$$\begin{aligned} M_1 &= \arg\max R^2_M \qquad (2) \\ &= \arg\min \left\{ \frac{2}{M}\mathbf{E}(Q_t) + (2\mathbf{E}(\theta_{t\varepsilon}) - 3\mathbf{E}(\sigma^4_{t\varepsilon}))M + M^2\mathbf{Var}(\sigma^2_{t\varepsilon}) \right\}, \qquad (3) \end{aligned}$$

where $Q_t = \int_t^{t+1} \sigma^4_s ds$, $\theta_{t\varepsilon} = \mathbf{E}(\varepsilon^4_t)$, *and* $\sigma^2_{t\varepsilon} = \mathbf{E}(\varepsilon^2_t) = 2\sigma^2_{tu}$.

Proof.

See Appendix.

Remark 1. (Interpretation.)

The optimal number of observations M_1 minimizes the unconditional variance of the regressor (realized variance). Under an assumption of independence between σ^2_{tu} and V_t (relaxed in Section 7), this minimization translates into maximization of the forecasting regression's R^2_M (as in ABM, 2006, and GS, 2006). The form of this unconditional variance is unusual and includes a term (of order M^2) which accounts for the variability of the noise variance (i.e., the last term in Eq. (2)).

Remark 2. (Estimation.)

The quantities $\theta_{t\varepsilon}$ and $\sigma^2_{t\varepsilon}$ can be estimated consistently (for each day in the sample) by using sample moments of the observed high-frequency return data (Bandi and Russell, 2003, 2006b).[1] Given $\theta_{t\varepsilon}$ and $\sigma^2_{t\varepsilon}$, consistent estimates of the unconditional moments $\mathbf{E}(\theta_{t\varepsilon})$, $\mathbf{E}(\sigma^4_{t\varepsilon})$, and $\mathbf{Var}(\sigma^2_{t\varepsilon})$ can be obtained under a stationarity assumption on the dynamic properties of the relevant series. Estimation of the daily quarticity

[1] Bandi and Russell (2006c) discuss finite sample bias corrections.

Q_t can be conducted by sampling observed returns at relatively low (15 or 20-minute) frequencies. (Bandi and Russell (2006a) discuss the empirical validity of this procedure by simulation.) Roughly unbiased estimates of the unconditional moment $\mathbf{E}(Q_t)$ can then be achieved by averaging the estimated daily quarticities under an assumption of stationarity for Q_t.

The Corollary below provides a simple, approximate rule to select M.

Corollary to Theorem 1.

Consider the regression in Eq. (1). *For a large optimal* M_1,

$$M_1^* \approx \arg\max R_M^2 = \left(\frac{\mathbf{E}(Q_t)}{\mathbf{Var}(\sigma_{t\varepsilon}^2)} \right)^{1/3}. \tag{4}$$

Remark 3

The approximate rule in Eq. (4) readily adapts to the noise variance's variance. The larger this variance is relative to the signal $\mathbf{E}(Q_t)$ coming from the underlying equilibrium price, the smaller the optimal number of observations needed to compute $\widehat{V}$. As always in these problems, a smaller number of observations translates into smaller noise contaminations.

Remark 4

The proposed approximate rule differs from the optimal (in an unconditional MSE sense), approximate rule proposed by Bandi and Russell (2003, 2006a) in the presence of time-invariant noise, i.e.,

$$M_2^* = \left(\frac{\mathbf{E}(Q_t)}{(\mathbf{E}(\varepsilon^2))^2} \right)^{1/3}. \tag{5}$$

It also differs from the optimal (in an R^2 sense), approximate rule proposed by ABM (2006) and GS (2006) in the case of time-invariant noise, i.e.,[2]

$$M_3^* = \left(\frac{2\mathbf{E}(Q_t)}{2\mathbf{E}(\varepsilon^4) - 3\mathbf{E}(\varepsilon^2)^2} \right)^{1/2}. \tag{6}$$

Naturally, the relative performance of these alternative rules depends on their relation with M_1^*. Importantly, empirically we find that $\mathbf{Var}(\sigma_{t\varepsilon}^2) > (\mathbf{E}(\varepsilon^2))^2 = (\mathbf{E}(\sigma_{t\varepsilon}^2))^2$. For our data, $M_3^* > M_2^* > M_1^*$ and, of course, $R^2_{M_1^*} > R^2_{M_2^*} > R^2_{M_3^*}$. This result is interesting. While a time-varying noise second moment can lead to relatively infrequent optimal sampling (M_1^*), optimizing the realized variance estimator's unconditional MSE (under the assumption of a constant noise variance), as implied by M_2^*, can be a superior strategy to focusing on the unconditional variance of realized variance (under the assumption of a constant bias term), as implied by M_3^*. As said, the latter would be the optimal choice, from a forecasting standpoint, should the second moment of the noise (or the realized variance estimator's bias) be assumed to be time-invariant.

[2] Interestingly, this is the same rule obtained by Bandi and Russell (2003) in a different context, namely the finite sample MSE (variance) minimization of their proposed bias-corrected realized variance estimator.

4 Conditional vs. Unconditional Frequency Choices

Rather than selecting one sampling frequency for all entries in the regressor vector (i.e., the vector of realized variance estimates), one could select a different optimal frequency for each entry/day. Bandi and Russell (2006a, 2006b) use this approach in predicting variance on the basis of autoregressive (fractionally-integrated) models. Whether the conditional approach has the potential to deliver superior forecasts than the unconditional approach described in the previous section depends on empirically verifiable conditions. Here, we consider the conditional MSE-based approximate rule in Bandi and Russell (2003, 2006a), i.e.,

$$M_{2t}^{*} = \left(\frac{Q_t}{\sigma_{t\varepsilon}^4}\right)^{1/3}, \tag{7}$$

and compare it to the optimal, approximate unconditional rule in Eq. (4).

Theorem 2.

Define

$$\begin{aligned}\mathbf{Var}_{M_1^*}(x_t) \quad = \quad & 2\left(\mathbf{Var}\left(\sigma_{t\varepsilon}^2\right)\right)^{1/3}\mathbf{E}(Q_t)^{2/3} + \left(\frac{\mathbf{E}(Q_t)}{\mathbf{Var}\left(\sigma_{t\varepsilon}^2\right)}\right)^{1/3}\mathbf{E}(2\theta_{t\varepsilon} - 3\sigma_{\varepsilon t}^4)\\ & + \left(4\mathbf{E}(\sigma_{t\varepsilon}^2 V_t) - \mathbf{E}(\theta_{t\varepsilon}) + 2\mathbf{E}\left(\sigma_{t\varepsilon}^4\right)\right)\\ & + \mathbf{Var}(V_t) + \left(\frac{\mathbf{E}(Q_t)}{\mathbf{Var}\left(\sigma_{t\varepsilon}^2\right)}\right)^{2/3}\left(\mathbf{Var}\left(\sigma_{t\varepsilon}^2\right)\right)\end{aligned}$$

and

$$\begin{aligned}\mathbf{Var}_{M_{2t}^*}(x_t) \quad = \quad & 2\mathbf{E}\left(\left(\sigma_{t\varepsilon}^2\right)^{2/3}\right)\mathbf{E}\left(Q_t^{2/3}\right) + \mathbf{E}\left(\left(\frac{Q_t}{\sigma_{\varepsilon t}^4}\right)^{1/3}(2\theta_{t\varepsilon} - 3\sigma_{t\varepsilon}^4)\right)\\ & + \left(4\mathbf{E}(\sigma_{t\varepsilon}^2 V_t) - \mathbf{E}(\theta_{t\varepsilon}) + 2\mathbf{E}\left(\sigma_{t\varepsilon}^4\right)\right)\\ & + \mathbf{Var}(V_t) + \mathbf{Var}\left((Q_t)^{1/3}\left(\sigma_{t\varepsilon}^2\right)^{1/3}\right) + 2\mathbf{Cov}\left(V_t, (Q_t)^{1/3}\left(\sigma_{t\varepsilon}^2\right)^{1/3}\right).\end{aligned}$$

If

$$\frac{\left(\mathbf{Var}(V_t) + \mathbf{Cov}\left(V_t, (Q_t)^{1/3}\left(\sigma_{t\varepsilon}^2\right)^{1/3}\right)\right)^2}{\mathbf{Var}_{M_{2t}^*}(x_t)} > \frac{(\mathbf{Var}(V_t))^2}{\mathbf{Var}_{M_1^*}(x_t)},$$

then

$$R^2_{M_{2t}^*} > R^2_{M_1^*}. \tag{8}$$

Proof.

See Appendix.

Remark 5

The statement in Theorem 2 highlights the moment condition affecting the preferability of an approximate conditional rule versus an approximate unconditional rule. In our sample, this moment condition is easily satisfied. Equivalently, we can provide a statement for exact conditional and unconditional rules. One could compare

$$\frac{\left(\mathbf{Var}(V_t)\right)^2}{\mathbf{Var}_{M_1}(x_t)} \tag{9}$$

with

$$\begin{aligned}\mathbf{Var}_{M_1}(x_t) &= \frac{2}{M_1}\mathbf{E}(Q_t) + (2\mathbf{E}(\theta_{t\varepsilon}) - 3\mathbf{E}(\sigma^4_{t\varepsilon}))M_1 \\ &+ \left(4\mathbf{E}(\sigma^2_{t\varepsilon}V_t) - \mathbf{E}(\theta_{t\varepsilon}) + 2\mathbf{E}\left(\sigma^4_{t\varepsilon}\right)\right) \\ &+\mathbf{Var}(V_t) + (M_1)^2\,\mathbf{Var}(\sigma^2_{t\varepsilon})\end{aligned}$$

to

$$\frac{\left(\mathbf{Var}(V_t) + \mathbf{Cov}\left(V_t, M_{2t}\sigma^2_{t\varepsilon}\right)\right)^2}{\mathbf{Var}_{M_{2t}}(x_t)} \tag{10}$$

with

$$\begin{aligned}\mathbf{Var}_{M_{2t}}(x_t) &= 2\mathbf{E}\left(\frac{Q_t}{M_{2t}}\right) + \mathbf{E}\left(M_{2t}(2\theta_{t\varepsilon} - 3\sigma^4_{t\varepsilon})\right) \\ &+ \left(4\mathbf{E}(\sigma^2_{t\varepsilon}V_t) - \mathbf{E}(\theta_{t\varepsilon}) + 2\mathbf{E}\left(\sigma^4_{t\varepsilon}\right)\right) \\ &+\mathbf{Var}(V_t) + \mathbf{Var}\left(M_{2t}\sigma^2_{t\varepsilon}\right) + 2\mathbf{Cov}\left(V_t, M_{2t}\sigma^2_{t\varepsilon}\right),\end{aligned}$$

where M_1 is the exact R^2-optimal number of observations (from Theorem 1) and M_{2t} is the exact (conditional) MSE-optimal number of observations from Bandi and Russell (2003, 2006a). If Eq. (10) is larger than Eq. (9), than $R^2_{M_{2t}} > R^2_{M_1}$. Naturally, this exact condition is slightly harder to verify than the useful, approximate condition in the theorem. Its verification requires solution of $n+1$, where n is the number of days in the sample, optimization problems to compute the relevant M's (i.e., M_1 and M_{2t} with $t = 1, ..., n$).

5 Forecasting Regressions

This section examines the implications of theory with data. We use SPIDERS (Standard and Poor's depository receipts) mid-quotes on the NYSE. We remove quotes whose associated price changes and/or spreads are larger than 10%.

In order to render the regressions feasible, we employ flat-top Bartlett kernels, flat-top cubic kernels, and flat-top modified Tukey-Hanning kernels, as proposed by Barndorff-Nielsen et al. (2006), to estimate V_{t+1}, i.e., the regressand. These estimators have favorable theoretical properties (unbiasedness and consistency) under IID noise and have been shown to perform well in practise (see, e.g., Barndorff-Nielsen et al., 2006, Bandi and Russell, 2005, and Bandi et al., 2006). Here we optimize them using methods discussed in Bandi and Russell (2005). Specifically, we select the number of autocovariances used for their computation in order to minimize their finite sample variance (conditionally, for each day in the sample).[3]

We run regressions of V_{t+1} on lagged realized variance. To fully capture the persistence properties of volatility, we run regressions of V_{t+1} on five lags of realized variance. We consider two subsets of data, the full period 1/2002 - 3/2006 and the shorter 1/2004 - 3/2006 period. In all cases, we use 1,000 observations to estimate the model's parameters and forecast. We report 6 cases for the regressors:

1. Realized variance with M_{2t}^*.
2. Realized variance with a conditional version of M_3^*, i.e., $M_{3t}^* = \left(\frac{2Q_t}{2\theta_{t\varepsilon} - 3\sigma_{t\varepsilon}^4}\right)^{1/2}$.
3. Realized variance with an unconditional choice of M obtained by least-squares minimization (jointly with estimation of the regression parameters), M_4^*.
4. Realized variance with $\overline{M}_{2t}^*$ (a unique M obtained by averaging the M_{2t}^*s across 1,000 observations - the same observations used for forecasting).
5. Realized variance with M_3^*.
6. Realized variance with $\overline{M}_{3t}^*$ (a unique M obtained by averaging the M_{3t}^*s across 1,000 observations - the same observations used for forecasting).

We begin by focusing on *unconditional* choices ((3) through (6)). The one-step estimator in (3) performs better than the estimator obtained by averaging conditional MSE-optimal frequencies in (4). In turn, the estimator in (4) performs better than the estimator (in (5)) obtained by choosing the unconditional optimal (in a variance sense) frequency. Notice that the one-step least-squares estimator is the estimator which *empirically* maximizes the regression's R^2 *jointly* with the regression parameters. In other words, this estimator represents the least-squares solution to frequency optimization in the context of forecasting. Under our assumptions, this estimator should provide a similar answer as the exact theoretical solution M_1. Hence, its superiority is not surprising. It is also not surprising that taking the time-varying noise into account

[3]Other estimators, such as the two-scale estimator of Zhang et al. (2005) and the multi-scale estimator of Zhang (2006), also have nice properties and could be used. We opt for the class of flat-top kernels because they have the same asymptotic properties and, under IID noise, have been shown to be unbiased in finite sample (Barndorff-Nielsen et al., 2006).

through an MSE criterion (as implied by (4)) appears to be a superior strategy than simply focusing on the unconditional variance of the regressor under the assumption of a time-invariant noise (as implied by (5)).

We find that $M_4^* < \overline{M}_{2t}^* < M_3^*$. In other words, $q_4^* > \overline{q}_{2t}^* > q_3^*$, where q indicates the average number of high-frequency observations to be skipped in the computation of the corresponding realized variance estimator. These values are reported in the tables. For the case of the flat-top Bartlett kernel in Table 1, for example, $q_4^* \approx 123$, $\overline{q}_{2t}^* \approx 74$, and $q_3^* = 55$. Since, empirically, $\mathbf{Var}(\sigma_{t\varepsilon}^2)(= 5.19e - 14) > \left(\mathbf{E}(\sigma_{t\varepsilon}^2)\right)^2 (= 1.6e - 14)$, the ranking of M values is consistent with theory. Differently put, it is consistent with the ranking reported in Section 3 above, namely $M_1^* < M_2^* < M_3^*$. This said, M_1^* and $\overline{M}_{2t}^*$ are not exacly equal to M_4^* and M_2^*, respectively. However, they are expected to be very close to them. While, as noted, M_4^* is the "empirical" R^2-maximizing number of observations, M_1^* is a useful approximation to it highlighting the main determinants of the optimal frequency in the time-varying noise case. In addition, $\overline{M}_{2t}^*$ and M_2^* are not identical objects due to Jensen's inequality-type arguments.

In our sample, averaging the conditional, optimal (in a variance sense) frequencies (as in (6)) is an inferior strategy to using the unconditional, variance-optimal frequency (as in (5)) advocated by ABM (2006) and GS (2006).

Turning to *conditional* choices, we find that using conditional MSE-optimal frequencies as in (1) and conditional variance-optimal frequencies as in (2) performs better than selecting only one frequency (the R^2-optimal frequency in (3)) for all realized variance entries. Importantly, as said earlier, the moment condition in Theorem 2 is easily satisfied in our sample.

These results are consistent across regressands and sample periods. Statistically, we find that (*i*) the forecast MSEs are always different from zero, (*ii*) a chi-squared test of the null hypothesis of equal MSEs across forecasting models is easily rejected, and (*iii*) pairwise t-tests of the null of equal MSEs between conditional and unconditional frequency choices are generally rejected in the low variance sample period 1/2004 - 3/2006 (Tables 4, 5, and 6).

6 Alternative Robust Regressors

We have shown that the presence of a time-varying bias should lead to careful consideration when forecasting volatility using realized variance. Not only is this type of bias not absorbed by the regression's intercept in general, but it can lead to unconditional choices of sampling frequency that are lower, rather than higher as in the constant bias case, than the realized variance estimator's unconditional MSE-optimal choice (under the assumption of a constant bias). We have also found that conditional choices can be beneficial in practise.

One could of course forecast variance using alternative, possibly more robust (to microstructure noise) regressors than the classical realized variance estimator. In this section we employ the same regressors that were used earlier as regressands, namely flat-top Bartlett kernels, flat-top cubic kernels, and flat-top modified

Tukey-Hanning kernels. When using the flat-top Bartlett kernel as a regressand, for instance, we also use it as a regressor. The forecasting set-up is the same as that in the previous section. We consider 5 lags for the regressor to capture volatility persistence effectively. We employ 1,000 observations to estimate the model's parameters and forecast. We now report 3 cases for the regressor:

1. Kernel estimates with a number of autocovariances chosen conditionally (for each day) in order to minimize the finite sample variance of the estimator as in Bandi and Russell (2005).

2. Kernel estimates with a number of autocovariances chosen unconditionally in order to minimize least-squares (i.e., jointly with the model's parameters).

3. Kernel estimates with a number of autocovariances chosen unconditionally as the average of the conditional variance-optimal choice in Bandi and Russell (2005).

As earlier, we start with *unconditional* choices. Even in the case of estimators with favorable (theoretical) bias properties in the presence of noise, there is scope for taking the documented time-variation in the noise second moment into account from a forecasting perspective. Unconditional autocovariance/bandwidth choices which empirically maximize the R^2 of the regression (as in the case of (2) above), and hence account for the presence of a time-varying bias directly, perform better than unconditional variance-optimal bandwidth choices (as in (3)). As thoroughly discussed by ABM (2006) and GS (2006), the latter would be optimal should the bias be constant (or absent, as is the case for the flat-top kernels in theory). It is interesting to notice that the number of autocovariances that is selected by the one-step R^2- optimal procedure is larger (around 13) then the number of autocovariances selected by the variance-optimal procedure (around 5). Qualitatively and quantitatively, this result is consistent with findings in Bandi and Russell (2005). There it was shown that the (conditional) finite sample optimization of the promising two-scale estimator results in a larger number of autocovariances than the finite sample optimization of the class of flat-top kernels proposed by Bardorff-Nielsen et al. (2006). This outcome was due to the presence of a finite sample bias in the case of the former and the need to reduce it in order to optimize the estimator's performance. Here, the divergence between the unconditional choices in (2) and (3) is possibly due to a time-varying bias component that is still left in the kernel estimates despite their favorable finite sample bias properties.

Turning to *conditional* bandwidth choices, choosing the autocovariances to optimize the finite sample variance of the estimator (as in (1)) for each entry in the regressor vector outperforms, in our sample, the "best" unconditional choice in (2). This is, again, consistent with our findings in the realized variance case.

Importantly, for our data the use of robust regressors leads, as expected, to gains over the use of the realized variance estimator. These gains, however, can be quite small in practise. Consider the case $V =$

flat-top Bartlett and $\widehat{V}$ = realized variance vs. the case V = flat-top Bartlett and $\widehat{V}$ = flat-top Bartlett, for the optimal one-step choice. The corresponding MSE values are $1.08e-09$ and $9.92e-10$. When using the flat-top cubic kernel, the values are $1.1819e-09$ and $1.1814e-09$, respectively. They are $1.0986e-09$ and $1.0753e-09$ in the case of the flat-top modified Tukey-Hanning kernel.

7 Extending the Framework: Dependence between Noise Variance and Equilibrium Price Variance

Sound economic reasoning suggests the potential for non-negligible cross-sectional dependencies between noise variance and equilibrium price variance. Higher uncertainty about the asset's equilibrium value (as represented by a higher V_t) implies higher likelihood of adverse price moves and hence higher inventory risk for the market maker. Similarly, higher uncertainty about the equilibrium value of the asset increases the market maker's risk of transacting with traders with superior information. Both risks ought to be compensated. Generally, other things equal, high equilibrium price variance translates into larger bid-ask spreads set by the market maker and hence larger noise variances. Bandi and Russell (2006b) provide discussions about the economic reasons underlying the *cross-sectional dependence* between equilibrium price variance and noise variance as well as empirical evidence for the S&P100 stocks.

A related issue has to do with the *time-series dependence* between equilibrium price variance and noise variance. From a forecasting perspective, if $\mathbf{Cov}(V_t, \sigma^2_{tu}) \neq 0$, the optimal R^2-rule to sample realized variance should be modified.

Theorem 3.

Consider the regression in Eq. (1). *If* $\mathbf{Cov}(V_t, \sigma^2_{tu}) \neq 0$, *then*

$$M_1^{\blacktriangle} = \arg\max R^2_M = \arg\max \frac{(\mathbf{Var}(V_t) + M\mathbf{Cov}(V_t, \sigma^2_{t\varepsilon}))^2}{\mathbf{Var}(x_t)}, \tag{11}$$

where

$$\begin{aligned}\mathbf{Var}(x_t) \quad &= \quad \frac{2}{M}\mathbf{E}(Q_t) + M\mathbf{E}(2\theta_{t\varepsilon} - 3\sigma^4_{t\varepsilon}) + \left(4\mathbf{E}(\sigma^2_{t\varepsilon}V_t) - \mathbf{E}(\theta_{t\varepsilon}) + 2\mathbf{E}\left(\sigma^4_{t\varepsilon}\right)\right) \\ &\quad\; + \mathbf{Var}\left(V_t\right) + M^2\mathbf{Var}\left(\sigma^2_{t\varepsilon}\right) + 2M\mathbf{Cov}(V_t, \sigma^2_{t\varepsilon}).\end{aligned}$$

Proof.

See Appendix.

Given consistent estimates of $\sigma^2_{t\varepsilon}$ for each day (obtained as described in Remark 2 above), the additional input $\mathbf{Cov}(V_t, \sigma^2_{t\varepsilon})$ can be readily calculated. Numerical maximization of R^2_M can then be conducted using standard methods.

Interestingly, if V_t and $\sigma^2_{t\varepsilon}$ are correlated in practise, the least-squares solution M_4^* provided above already accounts for a non-zero covariance between V_t and $\sigma^2_{t\varepsilon}$. This solution should of course be close to $M_1^{\blacktriangle}$. Having made this point, a thorough study of the dynamic properties of the bi-variate system $(V_t, \sigma^2_{t\varepsilon})$, and their implications for forecasting integrated variance, is an important topic of research better left for future work.

8 Conclusions

The second moment of market microstructure noise is time-varying. We provide a framework for analysing the impact of the noise variance's time variation in the context of linear volatility forecasting models.

In this framework, we find the need for lower sampling frequencies (in the case of realized variance) and larger bandwidth choices (in the case of kernel estimators of integrated variance) than required in linear volatility forecasting models in which microstructure noise is present but its variability is unaccounted for. We also show that frequency/bandwidth choices which adapt to the daily equilibrium price and noise features (named "conditional" in the text) have the potential to outperform (from an R^2 or forecast MSE standpoint) their unconditional counterparts.

9 Appendix

Proof of Theorem 1. Consider the d.g.p. $V_{t+1} = \alpha + \beta V_t + \varepsilon_{t+1}$, where ε_{t+1} is a forecast error uncorrelated with time t information. We run a regression of V_{t+1} on $x_t = V_t + U_t + (x_t - V_t - U_t) = V_t + U_t + \eta_t$, where $U_t = \mathbf{E}\left(\sum_{j=1}^{M} \varepsilon_j^2 | \sigma, \sigma_u\right) = M\sigma^2_{t\varepsilon}$, and $\eta_t = \left(\sum_{j=1}^{M} r_j^2 - V_t\right) + \sum_{j=1}^{M} r_j \varepsilon_j + \left(\sum_{j=1}^{M} \varepsilon_j^2 - U_t\right)$. Notice that $\mathbf{E}(\eta_t | \sigma, \sigma_u) = 0$. Now, consider $R^2 = \frac{\mathbf{Cov}^2(V_{t+1}, x_t)}{\mathbf{Var}(V_{t+1})\mathbf{Var}(x_t)}$. The (square root of the) numerator can be expressed as $\mathbf{Cov}\left(V_{t+1}, x_t\right) = \mathbf{Cov}(\alpha + \beta V_t + \varepsilon_{t+1}, V_t + U_t + \eta_t) = \beta \mathbf{Var}(V_t) + \beta \mathbf{Cov}(V_t, \eta_t)$ since $\mathbf{Cov}(V_t, U_t) = 0$. But,

$$
\begin{aligned}
\mathbf{Cov}(V_t, \eta_t) &= \mathbf{E}(V_t \eta_t) - \mathbf{E}(V_t)\mathbf{E}(\eta_t) \\
&= \mathbf{E}(V_t \eta_t) \\
&= \mathbf{E}\left(\mathbf{E}(V_t \eta_t | \sigma, \sigma_u)\right) \\
&= \mathbf{E}(V_t \mathbf{E}(\eta_t | \sigma, \sigma_u)) \\
&= 0.
\end{aligned}
$$

Hence, $\min_M \mathbf{Var}(x_t) \Rightarrow \max_M R^2$. By the law of total variance and Theorem 4 in Bandi and Russell (2006a), write

$$
\begin{aligned}
\mathbf{Var}(x_t) &= \mathbf{E}(\mathbf{Var}(x_t|\sigma, \sigma_u)) + \mathbf{Var}(\mathbf{E}(x_t|\sigma, \sigma_u)) \\
&= \mathbf{E}\left(\frac{2}{M} Q_t\right) + \mathbf{E}(M(2\theta_{t\varepsilon} - 3\sigma^4_{t\varepsilon})) + \left(4\mathbf{E}(\sigma^2_{t\varepsilon} V_t) - \mathbf{E}(\theta_{t\varepsilon}) + 2\mathbf{E}\left(\sigma^4_{t\varepsilon}\right)\right) \\
&\quad + \mathbf{Var}\left(V_t + M\sigma^2_{t\varepsilon}\right) \\
&= \frac{2}{M}\mathbf{E}(Q_t) + M\mathbf{E}(2\theta_{t\varepsilon} - 3\sigma^4_{t\varepsilon}) + \left(4\mathbf{E}(\sigma^2_{t\varepsilon} V_t) - \mathbf{E}(\theta_{t\varepsilon}) + 2\mathbf{E}\left(\sigma^4_{t\varepsilon}\right)\right) \\
&\quad + \mathbf{Var}\left(V_t\right) + M^2 \mathbf{Var}\left(\sigma^2_{t\varepsilon}\right). \qquad (12)
\end{aligned}
$$

■

Proof of Theorem 2. Consider Eq. (12). Plugging in the approximate unconditional rule $M_1^* = \left(\frac{\mathbf{E}(Q_t)}{\mathbf{Var}(\sigma_{t\varepsilon}^2)}\right)^{1/3}$ we obtain

$$\begin{aligned}
\mathbf{Var}_{M_1^*}(x_t) &= \frac{2}{\left(\frac{\mathbf{E}(Q_t)}{\mathbf{Var}(\sigma_{t\varepsilon}^2)}\right)^{1/3}}\mathbf{E}(Q_t) + \mathbf{E}\left(\left(\frac{\mathbf{E}(Q_t)}{\mathbf{Var}\left(\sigma_{t\varepsilon}^2\right)}\right)^{1/3}(2\theta_{t\varepsilon} - 3\sigma_{t\varepsilon}^4)\right) \\
&\quad + \left(4\mathbf{E}(\sigma_{t\varepsilon}^2 V_t) - \mathbf{E}(\theta_{t\varepsilon}) + 2\mathbf{E}\left(\sigma_{t\varepsilon}^4\right)\right) \\
&\quad + \mathbf{Var}\left(V_t\right) + \left(\frac{\mathbf{E}(Q_t)}{\mathbf{Var}\left(\sigma_{t\varepsilon}^2\right)}\right)^{2/3}\left(\mathbf{Var}\left(\sigma_{t\varepsilon}^2\right)\right) \\
&= 2\mathbf{Var}\left(\sigma_{t\varepsilon}^2\right)^{1/3}\mathbf{E}(Q_t)^{2/3} + \left(\frac{\mathbf{E}(Q_t)}{\mathbf{Var}\left(\sigma_{t\varepsilon}^2\right)}\right)^{1/3}\mathbf{E}(2\theta_{t\varepsilon} - 3\sigma_{t\varepsilon}^4) \\
&\quad + \left(4\mathbf{E}(\sigma_{t\varepsilon}^2 V_t) - \mathbf{E}(\theta_{t\varepsilon}) + 2\mathbf{E}\left(\sigma_{t\varepsilon}^4\right)\right) \\
&\quad + \mathbf{Var}\left(V_t\right) + \mathbf{E}(Q_t)^{2/3}\mathbf{Var}\left(\sigma_{t\varepsilon}^2\right)^{1/3}. \qquad (13)
\end{aligned}$$

Hence, $R^2_{M_1^*} = \frac{\beta^2(\mathbf{Var}(V_t))^2}{\mathbf{Var}(V_{t+1})\mathbf{Var}_{M_1^*}(x_t)}$. Using the conditional rule in Bandi and Russell (2003, 2006a):

$$\begin{aligned}
\mathbf{Var}_{M_{2t}^*}(x_t) = \mathbf{E}&\left(\frac{2}{\left(\frac{Q_t}{\sigma_{t\varepsilon}^4}\right)^{1/3}}Q_t\right) + \mathbf{E}\left(\left(\frac{Q_t}{\sigma_{t\varepsilon}^4}\right)^{1/3}(2\theta_{t\varepsilon} - 3\sigma_{t\varepsilon}^4)\right) \\
&+ \left(4\mathbf{E}(\sigma_{t\varepsilon}^2 V_t) - \mathbf{E}(\theta_{t\varepsilon}) + 2\mathbf{E}\left(\sigma_{t\varepsilon}^4\right)\right) + \mathbf{Var}\left(V_t + M\sigma_{t\varepsilon}^2\right) \\
= \ & 2\mathbf{E}\left(\left(\sigma_{t\varepsilon}^2\right)^{2/3}(Q_t)^{2/3}\right) + \mathbf{E}\left(\left(\frac{Q_t}{\sigma_{t\varepsilon}^4}\right)^{1/3}(2\theta_{t\varepsilon} - 3\sigma_{t\varepsilon}^4)\right) \\
&+ \left(4\mathbf{E}(\sigma_{t\varepsilon}^2 V_t) - \mathbf{E}(\theta_{t\varepsilon}) + 2\mathbf{E}\left(\sigma_{t\varepsilon}^4\right)\right) + \mathbf{Var}\left(V_t + \left(\frac{Q_t}{\sigma_{t\varepsilon}^4}\right)^{1/3}\sigma_{t\varepsilon}^2\right) \\
= \ & 2\mathbf{E}\left(\left(\sigma_{\varepsilon t}^2\right)^{2/3}\right)\mathbf{E}(Q_t^{2/3}) + \mathbf{E}\left(\left(\frac{Q_t}{\sigma_{t\varepsilon}^4}\right)^{1/3}(2\theta_{t\varepsilon} - 3\sigma_{t\varepsilon}^4)\right) \\
&+ \left(4\mathbf{E}(\sigma_{t\varepsilon}^2 V_t) - \mathbf{E}(\theta_{t\varepsilon}) + 2\mathbf{E}\left(\sigma_{t\varepsilon}^4\right)\right) \\
&+ \mathbf{Var}\left(V_t\right) + \mathbf{Var}\left((Q_t)^{1/3}\left(\sigma_{t\varepsilon}^2\right)^{1/3}\right) + 2\mathbf{Cov}(V_t, (Q_t)^{1/3}\left(\sigma_{t\varepsilon}^2\right)^{1/3}).
\end{aligned}$$

Thus, $R^2_{M_{2t}^*} = \frac{\beta^2(\mathbf{Var}(V_t)+\mathbf{Cov}(V_t,U_t))^2}{\mathbf{Var}(V_{t+1})\mathbf{Var}_{M_{2t}^*}(x_t)} = \frac{\beta^2\left(\mathbf{Var}(V_t)+\mathbf{Cov}\left(V_t,\left(\frac{Q_t}{\sigma_{t\varepsilon}^4}\right)^{1/3}\sigma_{t\varepsilon}^2\right)\right)^2}{\mathbf{Var}(V_{t+1})\mathbf{Var}_{M_{2t}^*}(x_t)} = \frac{\beta^2\left(\mathbf{Var}(V_t)+\mathbf{Cov}\left(V_t,(Q_t)^{1/3}\left(\sigma_{t\varepsilon}^2\right)^{1/3}\right)\right)^2}{\mathbf{Var}(V_{t+1})\mathbf{Var}_{M_{2t}^*}(x_t)}$.

■

Proof of Theorem 3. If $\mathbf{Cov}(V_t, \sigma_{tu}^2) \neq 0$, then $\mathbf{Cov}\left(V_{t+1}, x_t\right) = \mathbf{Cov}(\alpha + \beta V_t + \varepsilon_{t+1}, V_t + U_t + \eta_t) = \beta\mathbf{Var}(V_t) + \beta M\mathbf{Cov}(V_t, \sigma_{t\varepsilon}^2)$, from the proof of Theorem 1. In addition,

$$\begin{aligned}
\mathbf{Var}(x_t) &= \frac{2}{M}\mathbf{E}(Q_t) + \mathbf{E}(M(2\theta_{t\varepsilon} - 3\sigma_{t\varepsilon}^4)) + \left(4\mathbf{E}(\sigma_{t\varepsilon}^2 V_t) - \mathbf{E}(\theta_{t\varepsilon}) + 2\mathbf{E}\left(\sigma_{t\varepsilon}^4\right)\right) \\
&\quad + \mathbf{Var}\left(V_t\right) + M^2\mathbf{Var}\left(\sigma_{t\varepsilon}^2\right) + 2M\mathbf{Cov}(V_t, \sigma_{t\varepsilon}^2). \qquad (14)
\end{aligned}$$

Hence,

$$R^2 = \frac{\beta^2}{\mathbf{Var}(V_{t+1})}\frac{(\mathbf{Var}(V_t) + M\mathbf{Cov}(V_t, \sigma_{\varepsilon t}^2))^2}{\mathbf{Var}(x_t)}$$

with $\mathbf{Var}(x_t)$ defined in Eq. (14).

■

References

[1] Aït-Sahalia, Y., and L. Mancini (2006). Out-of-sample forecasts of quadratic variation. *Working paper.*

[2] Andersen, T.G., T. Bollerslev, and N. Meddahi (2004). Analytic evaluation of volatility forecasts. *International Economic Review 45*, 1079-1110.

[3] Andersen, T.G., T. Bollerslev, and N. Meddahi (2006). Realized volatility forecasting and market microstructure noise. *Working paper.*

[4] Andersen, T.G., T. Bollerslev, F.X. Diebold, and P. Labys (2003). Modeling and forecasting realized volatility. *Econometrica 71*, 579-625.

[5] Bandi, F.M., and J.R. Russell (2003). Microstructure noise, realized volatility, and optimal sampling. *Working paper.*

[6] Bandi, F.M., and J.R. Russell (2005). Market microstructure noise, integrated variance estimators, and the accuracy of asymptotic approximations. *Working paper.*

[7] Bandi, F.M., and J.R. Russell (2006a). Microstructure noise, realized variance, and optimal sampling. *Working paper.*

[8] Bandi, F.M., and J.R. Russell (2006b). Separating microstructure noise from volatility. *Journal of Financial Economics 79,* 655-692.

[9] Bandi, F.M., and J.R. Russell (2006c). Volatility. In J. Birge and V. Linetski (Eds.) *Handbook of financial engineering. Elsevier, North Holland,* forthcoming.

[10] Bandi, F.M., J.R. Russell, and C. Yang (2006). Realized volatility forecasting and option pricing. *Working paper.*

[11] Barndorff-Nielsen, O.E., and N. Shephard (2002). Econometric analysis of realized volatility and its use in estimating stochastic volatility models. *Journal of the Royal Statistical Society, Series B 64*, 253-280.

[12] Barndorff-Nielsen, O.E., and N. Shephard (2006). Variation, jumps, market frictions and high-frequency data in financial econometrics. In R. Blundell, P. Torsten and W. K. Newey (Eds.) *Advances in Economics and Econometrics. Theory and Applications.* Ninth World Congress. Econometric Society Monographs, Cambridge University Press, forthcoming.

[13] Barndorff-Nielsen, O.E., P. Hansen, A. Lunde and N. Shephard (2006). Designing realized kernels to measure ex-post variation of equity prices in the presence of noise. *Working paper.*

[14] Corradi, V., W. Distaso, and N. Swanson (2006). Predictive inference for integrated volatility. *Working paper.*

[15] Ghysels, E. and A. Sinko (2006a). Comment on Hansen and Lunde. *Journal of Business and Economic Statistics 24,* 192-194.

[16] Ghysels, E. and A. Sinko (2006b). Volatility forecasting and microstructure noise. *Working paper.*

[17] Hansen, P.R., and A. Lunde (2006). Realized variance and market microstructure noise (with discussions). *Journal of Business and Economic Statistics 24,* 127-161.

[18] McAleer, M., and M. Medeiros (2006). Realized volatility: A review. *Working paper.*

[19] Oomen, R.C.A. (2006). Properties of realized variance under alternative sampling schemes. *Journal of Business and Economic Statistics 24,* 219-237.

[20] Zhang, L., P. Mykland, and Y. Aït-Sahalia (2005). A tale of two time scales: determining integrated volatility with noisy high-frequency data. *Journal of the American Statistical Association 100,* 1394-1411.

[21] Zhang, L. (2006). Efficient estimation of stochastic volatility using noisy observations: a multi-scale approach. *Bernoulli, forthcoming.*

[22] Zhou, B. (1996). High-frequency data and volatility in foreign-exchange rates. *Journal of Business and Economic Statistics 14*, 45-52.

Figures and Tables

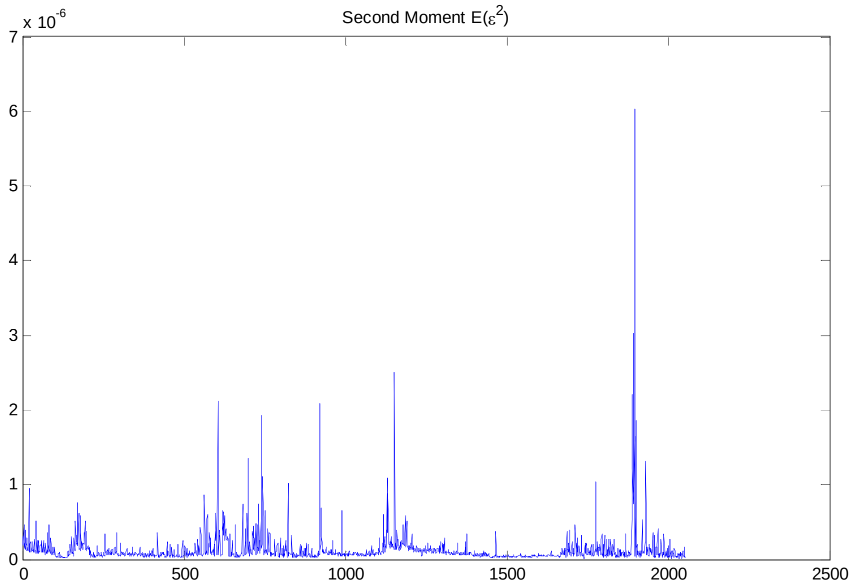

Figure 1. Microstructure noise second moment's estimates. We use Spiders mid-quotes on the NYSE. The sample period is 2002/1 - 2006/3.

LHS = FlatBart **RHS = RV** **5 lags**

2002/1-2006/3

(1)	Avg. q	MSE	Mean(Vtrue)	Mean(Vfore)	Std(Vtrue)	Std(Vfore)
	162.326	**9.66E-10**	4.67E-05	5.88E-05	5.00E-05	4.93E-05
	Avg. a0	Avg. a1	Avg. a2	Avg. a3	Avg. a4	Avg. a5
	2.32E-05	0.281312	0.0989931	-0.0033682	0.0937042	0.045113
(2)	Avg. q	MSE	Mean(Vtrue)	Mean(Vfore)	Std(Vtrue)	Std(Vfore)
	117.478	**9.98E-10**	4.67E-05	5.80E-05	5.00E-05	4.29E-05
	Avg. a0	Avg. a1	Avg. a2	Avg. a3	Avg. a4	Avg. a5
	2.43E-05	0.318263	0.0556906	0.0192768	0.102851	0.00968516
(3)	Avg. q	MSE	Mean(Vtrue)	Mean(Vfore)	Std(Vtrue)	Std(Vfore)
	123.879	**1.08E-09**	4.67E-05	6.26E-05	5.00E-05	4.86E-05
	Avg. a0	Avg. a1	Avg. a2	Avg. a3	Avg. a4	Avg. a5
	2.77E-05	0.219259	0.0998483	0.0352042	0.0524899	0.0425611
(4)	Avg. q	MSE	Mean(Vtrue)	Mean(Vfore)	Std(Vtrue)	Std(Vfore)
	74.8889	**1.09E-09**	4.67E-05	6.18E-05	5.00E-05	4.51E-05
	Avg. a0	Avg. a1	Avg. a2	Avg. a3	Avg. a4	Avg. a5
	2.94E-05	0.210582	0.0894445	0.0269597	0.0575391	0.0497621
(5)	Avg. q	MSE	Mean(Vtrue)	Mean(Vfore)	Std(Vtrue)	Std(Vfore)
	55.8889	**1.13E-09**	4.67E-05	6.25E-05	5.00E-05	4.57E-05
	Avg. a0	Avg. a1	Avg. a2	Avg. a3	Avg. a4	Avg. a5
	2.86E-05	0.212029	0.0925057	0.0296564	0.055968	0.0469369
(6)	Avg. q	MSE	Mean(Vtrue)	Mean(Vfore)	Std(Vtrue)	Std(Vfore)
	53.9468	**1.19E-09**	4.67E-05	6.26E-05	5.00E-05	4.54E-05
	Avg. a0	Avg. a1	Avg. a2	Avg. a3	Avg. a4	Avg. a5
	2.79E-05	0.21654	0.0936616	0.0285354	0.0555054	0.0496944

	(1)	(2)	(3)	(4)	(5)	(6)
MSE	9.66E-10	9.98E-10	1.08E-09	1.09E-09	1.13E-09	1.19E-09
HAC_std	2.34E-10	2.41E-10	2.20E-10	2.25E-10	2.10E-10	2.40E-10
t_MSE	4.1291	4.13497	4.91409	4.85042	5.37764	4.97761
	t12	t13	t14	t15	t25	
T	-0.24062	-1.32543	-1.54206	-1.66828	-1.02195	

Joint Chi-squared test: 50.41

Table 1. We report forecasting regressions of integrated variance (estimated using optimally-defined flat-top Bartlett kernels as in Bandi and Russell, 2005) on five lags of realized variance. The regressor (realized variance) is sampled using 6 methods described in the main text. We use Spiders mid-quotes on the NYSE. The sample period is 2002/1 - 2006/3. In all cases, 1,000 observations are employed to estimate the model parameters (a0 through a5) and forecast. The table reports means and standard deviations for both the regressand (true) and the volatility forecasts. It also reports the average number of observations to be skipped (q) according to each sampling rule. T-statistics for the individual MSEs, t-statistics for pair-wise tests of equal MSEs, and a joint Chi-squared test of equal MSEs across sampling methods are reported in the second panel.

LHS = FlatCubic RHS = RV 5 lags
2002/1-2006/3

(1)	Avg. q	MSE	Mean(Vtrue)	Mean(Vfore)	Std(Vtrue)	Std(Vfore)
	162.326	**1.05E-09**	4.85E-05	5.98E-05	5.19E-05	4.87E-05
	Avg. a0	Avg. a1	Avg. a2	Avg. a3	Avg. a4	Avg. a5
	2.42E-05	0.287751	0.10447	-0.0050108	0.0934314	0.0368125
(2)	Avg. q	MSE	Mean(Vtrue)	Mean(Vfore)	Std(Vtrue)	Std(Vfore)
	117.478	**1.07E-09**	4.85E-05	5.90E-05	5.19E-05	4.23E-05
	Avg. a0	Avg. a1	Avg. a2	Avg. a3	Avg. a4	Avg. a5
	2.50E-05	0.324704	0.0627989	0.0173676	0.102172	0.00247085
(3)	Avg. q	MSE	Mean(Vtrue)	Mean(Vfore)	Std(Vtrue)	Std(Vfore)
	117.969	**1.1819E-09**	4.85E-05	6.41E-05	5.19E-05	4.82E-05
	Avg. a0	Avg. a1	Avg. a2	Avg. a3	Avg. a4	Avg. a5
	2.86E-05	0.222019	0.103153	0.0380231	0.0497926	0.0371127
(4)	Avg. q	MSE	Mean(Vtrue)	Mean(Vfore)	Std(Vtrue)	Std(Vfore)
	74.8889	**1.1871E-09**	4.85E-05	6.29E-05	5.19E-05	4.46E-05
	Avg. a0	Avg. a1	Avg. a2	Avg. a3	Avg. a4	Avg. a5
	3.03E-05	0.21586	0.0941549	0.0257633	0.0545427	0.0457488
(5)	Avg. q	MSE	Mean(Vtrue)	Mean(Vfore)	Std(Vtrue)	Std(Vfore)
	55.8889	**1.22E-09**	4.85E-05	6.36E-05	5.19E-05	4.52E-05
	Avg. a0	Avg. a1	Avg. a2	Avg. a3	Avg. a4	Avg. a5
	2.95E-05	0.217557	0.0973909	0.0281942	0.0531524	0.0425865
(6)	Avg. q	MSE	Mean(Vtrue)	Mean(Vfore)	Std(Vtrue)	Std(Vfore)
	53.9468	**1.30E-09**	4.85E-05	6.37E-05	5.19E-05	4.50E-05
	Avg. a0	Avg. a1	Avg. a2	Avg. a3	Avg. a4	Avg. a5
	2.88E-05	0.222269	0.0983164	0.0269882	0.0528116	0.0453233

	(1)	(2)	(3)	(4)	(5)	(6)
MSE	1.05E-09	1.07E-09	1.1819E-09	1.1871E-09	1.22E-09	1.30E-09
HAC_std	2.78E-10	2.84E-10	2.60E-10	2.66E-10	2.50E-10	2.85E-10
t_MSE	3.77433	3.77673	4.57811	4.46129	4.88257	4.5744
	t12	t13	t14	t15	t25	
T	-0.1656	-1.4041	-1.50741	-1.57473	-1.01939	

Joint Chi-squared test: 43.78

Table 2. We report forecasting regressions of integrated variance (estimated using optimally-defined flat-top cubic kernels as in Bandi and Russell, 2005) on five lags of realized variance. The regressor (realized variance) is sampled using 6 methods described in the main text. We use Spiders mid-quotes on the NYSE. The sample period is 2002/1 - 2006/3. In all cases, 1,000 observations are employed to estimate the model parameters (a0 through a5) and forecast. The table reports means and standard deviations for both the regressand (true) and the volatility forecasts. It also reports the average number of observations to be skipped (q) according to each sampling rule. T-statistics for the individual MSEs, t-statistics for pair-wise tests of equal MSEs, and a joint Chi-squared test of equal MSEs across sampling methods are reported in the second panel.

LHS = FlatTukey RHS = RV 5 lags
2002/1-2006/3

(1)	Avg. q	MSE	Mean(Vtrue)	Mean(Vfore)	Std(Vtrue)	Std(Vfore)
	162.326	**9.66E-10**	4.81E-05	5.94E-05	5.10E-05	4.90E-05
	Avg. a0	Avg. a1	Avg. a2	Avg. a3	Avg. a4	Avg. a5
	2.38E-05	0.283859	0.10733	-0.0055169	0.0885941	0.0442513
(2)	Avg. q	MSE	Mean(Vtrue)	Mean(Vfore)	Std(Vtrue)	Std(Vfore)
	117.478	**1.02E-09**	4.81E-05	5.87E-05	5.10E-05	4.25E-05
	Avg. a0	Avg. a1	Avg. a2	Avg. a3	Avg. a4	Avg. a5
	2.47E-05	0.321145	0.0632796	0.0151823	0.0997437	0.00968388
(3)	Avg. q	MSE	Mean(Vtrue)	Mean(Vfore)	Std(Vtrue)	Std(Vfore)
	122.966	**1.0986E-09**	4.81E-05	6.35E-05	5.10E-05	4.85E-05
	Avg. a0	Avg. a1	Avg. a2	Avg. a3	Avg. a4	Avg. a5
	2.82E-05	0.221173	0.10493	0.0344397	0.049225	0.0417724
(4)	Avg. q	MSE	Mean(Vtrue)	Mean(Vfore)	Std(Vtrue)	Std(Vfore)
	74.8889	**1.1050E-09**	4.81E-05	6.25E-05	5.10E-05	4.48E-05
	Avg. a0	Avg. a1	Avg. a2	Avg. a3	Avg. a4	Avg. a5
	2.99E-05	0.212318	0.0955053	0.0245537	0.0530928	0.0510944
(5)	Avg. q	MSE	Mean(Vtrue)	Mean(Vfore)	Std(Vtrue)	Std(Vfore)
	55.8889	**1.13E-09**	4.81E-05	6.32E-05	5.10E-05	4.54E-05
	Avg. a0	Avg. a1	Avg. a2	Avg. a3	Avg. a4	Avg. a5
	2.91E-05	0.214138	0.0991262	0.0267773	0.0520487	0.0472941
(6)	Avg. q	MSE	Mean(Vtrue)	Mean(Vfore)	Std(Vtrue)	Std(Vfore)
	53.9468	**1.22E-09**	4.81E-05	6.33E-05	5.10E-05	4.52E-05
	Avg. a0	Avg. a1	Avg. a2	Avg. a3	Avg. a4	Avg. a5
	2.84E-05	0.218807	0.0997764	0.0255501	0.0515927	0.0504963

	(1)	(2)	(3)	(4)	(5)	(6)
MSE	9.66E-10	1.02E-09	1.0986E-09	1.1050E-09	1.13E-09	1.22E-09
HAC_std	2.43E-10	2.59E-10	2.33E-10	2.37E-10	2.20E-10	2.56E-10
t_MSE	3.96744	3.94106	4.72397	4.66092	5.15286	4.74954
	t12	T13	t14	t15	t25	
T	-0.39802	-1.40348	-1.62466	-1.70449	-0.848255	

Joint Chi-squared test: 48.32

Table 3. We report forecasting regressions of integrated variance (estimated using optimally-defined flat-top modified Tukey-Hanning kernels as in Bandi and Russell, 2005) on five lags of realized variance. The regressor (realized variance) is sampled using 6 methods described in the main text. We use Spiders mid-quotes on the NYSE. The sample period is 2002/1 - 2006/3. In all cases, 1,000 observations are employed to estimate the model parameters (a0 through a5) and forecast. The table reports means and standard deviations for both the regressand (true) and the volatility forecasts. It also reports the average number of observations to be skipped (q) according to each sampling rule. T-statistics for the individual MSEs, t-statistics for pair-wise tests of equal MSEs, and a joint Chi-squared test of equal MSEs across sampling methods are reported in the second panel.

LHS = FlatBart RHS = RV 5 lags
2004/1-2006/3

(1)	Avg. q	MSE	Mean(Vtrue)	Mean(Vfore)	Std(Vtrue)	Std(Vfore)
	171.112	**2.35E-10**	2.82E-05	3.20E-05	1.52E-05	9.51E-06
	Avg. a0	Avg. a1	Avg. a2	Avg. a3	Avg. a4	Avg. a5
	1.49E-05	0.276181	0.0922763	0.0236755	0.0942341	0.0224669
(2)	Avg. q	MSE	Mean(Vtrue)	Mean(Vfore)	Std(Vtrue)	Std(Vfore)
	178.018	**2.47E-10**	2.82E-05	3.12E-05	1.52E-05	1.11E-05
	Avg. a0	Avg. a1	Avg. a2	Avg. a3	Avg. a4	Avg. a5
	1.28E-05	0.342192	0.0666968	0.00278203	0.126232	-0.002328
(3)	Avg. q	MSE	Mean(Vtrue)	Mean(Vfore)	Std(Vtrue)	Std(Vfore)
	160.244	**4.11E-10**	2.82E-05	3.81E-05	1.52E-05	1.48E-05
	Avg. a0	Avg. a1	Avg. a2	Avg. a3	Avg. a4	Avg. a5
	2.10E-05	0.187816	0.0960295	0.0664505	0.0519591	0.021378
(4)	Avg. q	MSE	Mean(Vtrue)	Mean(Vfore)	Std(Vtrue)	Std(Vfore)
	65.9439	**5.36E-10**	2.82E-05	3.87E-05	1.52E-05	1.88E-05
	Avg. a0	Avg. a1	Avg. a2	Avg. a3	Avg. a4	Avg. a5
	1.84E-05	0.224782	0.0967559	0.0235081	0.0666108	0.0378135
(5)	Avg. q	MSE	Mean(Vtrue)	Mean(Vfore)	Std(Vtrue)	Std(Vfore)
	71.2857	**5.57E-10**	2.82E-05	3.88E-05	1.52E-05	1.91E-05
	Avg. a0	Avg. a1	Avg. a2	Avg. a3	Avg. a4	Avg. a5
	1.97E-05	0.220002	0.0931147	0.0254188	0.0662948	0.0317596
(6)	Avg. q	MSE	Mean(Vtrue)	Mean(Vfore)	Std(Vtrue)	Std(Vfore)
	69.6673	**6.04E-10**	2.82E-05	3.91E-05	1.52E-05	2.06E-05
	Avg. a0	Avg. a1	Avg. a2	Avg. a3	Avg. a4	Avg. a5
	1.87E-05	0.225986	0.0959497	0.0224315	0.0649285	0.0383145

	(1)	(2)	(3)	(4)	(5)	(6)
MSE	2.35E-10	2.47E-10	4.11E-10	5.36E-10	5.57E-10	6.04E-10
HAC_std	2.20E-11	2.52E-11	7.90E-11	1.48E-10	1.49E-10	2.04E-10
t_MSE	10.674	9.80144	5.20368	3.6313	3.73351	2.95806
	t12	t13	t14	t15	t25	
T	-1.54609	-2.38177	-2.09481	-2.21819	-2.11737	

Joint Chi-square test: 117.82

Table 4. We report forecasting regressions of integrated variance (estimated using optimally-sampled flat-top Bartlett kernels as in Bandi and Russell, 2005) on five lags of realized variance. The regressor (realized variance) is sampled using 6 methods described in the main text. We use Spiders mid-quotes on the NYSE. The sample period is 2004/1 - 2006/3. In all cases, 1,000 observations are employed to estimate the model parameters (a0 through a5) and forecast. The table reports means and standard deviations for both the regressand (true) and the volatility forecasts. It also reports the average number of observations to be skipped (q) according to each sampling rule. T-statistics for the individual MSEs, t-statistics for pair-wise tests of equal MSEs, and a joint Chi-squared test of equal MSEs across sampling methods are reported in the second panel.

LHS = FlatCubic RHS = RV 5 lags

2004/1-2006/3

(1)	Avg. q	MSE	Mean(Vtrue)	Mean(Vfore)	Std(Vtrue)	Std(Vfore)
	171.112	**2.31E-10**	2.94E-05	3.36E-05	1.56E-05	9.47E-06
	Avg. a0	Avg. a1	Avg. a2	Avg. a3	Avg. a4	Avg. a5
	1.63E-05	0.285943	0.100449	0.0191892	0.0940962	0.012773
(2)	Avg. q	MSE	Mean(Vtrue)	Mean(Vfore)	Std(Vtrue)	Std(Vfore)
	178.018	**2.36E-10**	2.94E-05	3.27E-05	1.56E-05	1.11E-05
	Avg. a0	Avg. a1	Avg. a2	Avg. a3	Avg. a4	Avg. a5
	1.42E-05	0.351297	0.0773286	-0.00088397	0.125422	-0.0128982
(3)	Avg. q	MSE	Mean(Vtrue)	Mean(Vfore)	Std(Vtrue)	Std(Vfore)
	145.897	**4.90E-10**	2.94E-05	4.08E-05	1.56E-05	1.68E-05
	Avg. a0	Avg. a1	Avg. a2	Avg. a3	Avg. a4	Avg. a5
	2.27E-05	0.188722	0.100167	0.0738194	0.047709	0.0128716
(4)	Avg. q	MSE	Mean(Vtrue)	Mean(Vfore)	Std(Vtrue)	Std(Vfore)
	65.9439	**5.65E-10**	2.94E-05	4.04E-05	1.56E-05	1.91E-05
	Avg. a0	Avg. a1	Avg. a2	Avg. a3	Avg. a4	Avg. a5
	2.00E-05	0.232728	0.104053	0.0209948	0.0620747	0.0318149
(5)	Avg. q	MSE	Mean(Vtrue)	Mean(Vfore)	Std(Vtrue)	Std(Vfore)
	71.2857	**5.89E-10**	2.94E-05	4.05E-05	1.56E-05	1.94E-05
	Avg. a0	Avg. a1	Avg. a2	Avg. a3	Avg. a4	Avg. a5
	2.13E-05	0.227631	0.10084	0.0228622	0.0618063	0.0257183
(6)	Avg. q	MSE	Mean(Vtrue)	Mean(Vfore)	Std(Vtrue)	Std(Vfore)
	69.6673	**6.42E-10**	2.94E-05	4.08E-05	1.56E-05	2.11E-05
	Avg. a0	Avg. a1	Avg. a2	Avg. a3	Avg. a4	Avg. a5
	2.03E-05	0.234069	0.103272	0.0199102	0.0602224	0.0322606

	(1)	(2)	(3)	(4)	(5)	(6)
MSE	2.31E-10	2.36E-10	4.90E-10	5.65E-10	5.89E-10	6.42E-10
HAC_std	2.31E-11	2.56E-11	1.15E-10	1.68E-10	1.69E-10	2.28E-10
t_MSE	9.97592	9.21065	4.25107	3.36381	3.48204	2.81823
	t12	t13	t14	t15	t25	
T	-0.73573	-2.31294	-2.01153	-2.14223	-2.10502	

Joint Chi-squared test: 105.71

Table 5. We report forecasting regressions of integrated variance (estimated using optimally-defined flat-top cubic kernels as in Bandi and Russell, 2005) on five lags of realized variance. The regressor (realized variance) is sampled using 6 methods described in the main text. We use Spiders mid-quotes on the NYSE. The sample period is 2004/1 - 2006/3. In all cases, 1,000 observations are employed to estimate the model parameters (a0 through a5) and forecast. The table reports means and standard deviations for both the regressand (true) and the volatility forecasts. It also reports the average number of observations to be skipped (q) according to each sampling rule. T-statistics for the individual MSEs, t-statistics for pair-wise tests of equal MSEs, and a joint Chi-squared test of equal MSEs across sampling methods are reported in the second panel.

LHS = FlatTukey RHS = RV 5 lags
2004/1-2006/3

(1)	Avg. q	MSE	Mean(Vtrue)	Mean(Vfore)	Std(Vtrue)	Std(Vfore)
	171.112	**2.34E-10**	2.94E-05	3.30E-05	1.57E-05	9.43E-06
	Avg. a0	Avg. a1	Avg. a2	Avg. a3	Avg. a4	Avg. a5
	1.58E-05	0.279715	0.103455	0.0165643	0.0901693	0.0233784
(2)	Avg. q	MSE	Mean(Vtrue)	Mean(Vfore)	Std(Vtrue)	Std(Vfore)
	178.018	**2.42E-10**	2.94E-05	3.23E-05	1.57E-05	1.11E-05
	Avg. a0	Avg. a1	Avg. a2	Avg. a3	Avg. a4	Avg. a5
	1.38E-05	0.346032	0.07832	-0.0041256	0.121751	-0.0022715
(3)	Avg. q	MSE	Mean(Vtrue)	Mean(Vfore)	Std(Vtrue)	Std(Vfore)
	158.505	**4.55E-10**	2.94E-05	3.96E-05	1.57E-05	1.64E-05
	Avg. a0	Avg. a1	Avg. a2	Avg. a3	Avg. a4	Avg. a5
	2.20E-05	0.189851	0.103888	0.0649971	0.0479956	0.0198157
(4)	Avg. q	MSE	Mean(Vtrue)	Mean(Vfore)	Std(Vtrue)	Std(Vfore)
	65.9439	**5.50E-10**	2.94E-05	3.99E-05	1.57E-05	1.90E-05
	Avg. a0	Avg. a1	Avg. a2	Avg. a3	Avg. a4	Avg. a5
	1.94E-05	0.227948	0.105575	0.0190842	0.0605393	0.0394834
(5)	Avg. q	MSE	Mean(Vtrue)	Mean(Vfore)	Std(Vtrue)	Std(Vfore)
	71.2857	**5.76E-10**	2.94E-05	3.99E-05	1.57E-05	1.93E-05
	Avg. a0	Avg. a1	Avg. a2	Avg. a3	Avg. a4	Avg. a5
	2.07E-05	0.222864	0.103168	0.0205479	0.0605758	0.0326004
(6)	Avg. q	MSE	Mean(Vtrue)	Mean(Vfore)	Std(Vtrue)	Std(Vfore)
	69.6673	**6.21E-10**	2.94E-05	4.03E-05	1.57E-05	2.09E-05
	Avg. a0	Avg. a1	Avg. a2	Avg. a3	Avg. a4	Avg. a5
	1.97E-05	0.228984	0.105312	0.0172791	0.0590676	0.0400023

	(1)	(2)	(3)	(4)	(5)	(6)
MSE	2.34E-10	2.42E-10	4.55E-10	5.50E-10	5.76E-10	6.21E-10
HAC_std	2.45E-11	2.71E-11	1.06E-10	1.65E-10	1.66E-10	2.21E-10
t_MSE	9.54519	8.91172	4.3104	3.33834	3.46247	2.80679
	t12	t13	t14	t15	t25	
T	-0.97995	-2.21054	-1.94983	-2.08929	-2.03753	

Joint Chi-squared test: 95.5

Table 6. We report forecasting regressions of integrated variance (estimated using optimally-defined flat-top modified Tukey-Hanning kernels as in Bandi and Russell, 2005) on five lags of realized variance. The regressor (realized variance) is sampled using 6 methods described in the main text. We use Spiders mid-quotes on the NYSE. The sample period is 2004/1 - 2006/3. In all cases, 1,000 observations are employed to estimate the model parameters (a0 through a5) and forecast. The table reports means and standard deviations for both the regressand (true) and the volatility forecasts. It also reports the average number of observations to be skipped (q) according to each sampling rule. T-statistics for the individual MSEs, t-statistics for pair-wise tests of equal MSEs, and a joint Chi-squared test of equal MSEs across sampling methods are reported in the second panel.

LHS = FlatBart RHS = FlatBart 5 lags
2002/1-2006/3

(1)	Avg. q	MSE	Mean(Vtrue)	Mean(Vfore)	Std(Vtrue)	Std(Vfore)
	10.2688	**9.00E-10**	4.67E-05	5.38E-05	5.00E-05	3.61E-05
	Avg. a0	Avg. a1	Avg. a2	Avg. a3	Avg. a4	Avg. a5
	1.94E-05	0.364522	0.141937	0.0872028	0.099541	0.0647708
(2)	Avg. q	MSE	Mean(Vtrue)	Mean(Vfore)	Std(Vtrue)	Std(Vfore)
	13.3276	**9.92E-10**	4.67E-05	4.99E-05	5.00E-05	3.17E-05
	Avg. a0	Avg. a1	Avg. a2	Avg. a3	Avg. a4	Avg. a5
	2.28E-05	0.307943	0.111025	0.0200002	0.091199	0.0587348
(3)	Avg. q	MSE	Mean(Vtrue)	Mean(Vfore)	Std(Vtrue)	Std(Vfore)
	5.49193	**1.03E-09**	4.67E-05	5.37E-05	5.00E-05	3.48E-05
	Avg. a0	Avg. a1	Avg. a2	Avg. a3	Avg. a4	Avg. a5
	2.27E-05	0.332203	0.125932	0.0373003	0.09843	0.069047

Table 7. We report forecasting regressions of integrated variance (estimated using optimally-defined flat-top Bartlett kernels as in Bandi and Russell, 2005) on five lags of Bartlett kernel estimates. The regressor's autocovariances are chosen using 3 methods described in the main text. We use Spiders mid-quotes on the NYSE. The sample period is 2002/1 - 2006/3. In all cases, 1,000 observations are employed to estimate the model parameters (a0 through a5) and forecast. The table reports means and standard deviations for both the regressand (true) and the volatility forecasts. It also reports the average number of autocovariances selected according to each rule.

LHS = FlatCubic **RHS = FlatCubic** **5 lags**
2002/1-2006/3

(1)	Avg. q	MSE	Mean(Vtrue)	Mean(Vfore)	Std(Vtrue)	Std(Vfore)
	8.83476	**1.02E-09**	4.85E-05	5.55E-05	5.19E-05	3.67E-05
	Avg. a0	Avg. a1	Avg. a2	Avg. a3	Avg. a4	Avg. a5
	2.02E-05	0.367511	0.144537	0.0827307	0.096981	0.0580071
(2)	Avg. q	MSE	Mean(Vtrue)	Mean(Vfore)	Std(Vtrue)	Std(Vfore)
	13.0978	**1.1814E-09**	4.85E-05	5.18E-05	5.19E-05	3.59E-05
	Avg. a0	Avg. a1	Avg. a2	Avg. a3	Avg. a4	Avg. a5
	2.39E-05	0.308348	0.112293	0.0176941	0.090176	0.045786
(3)	Avg. q	MSE	Mean(Vtrue)	Mean(Vfore)	Std(Vtrue)	Std(Vfore)
	5.03989	**1.1908E-09**	4.85E-05	5.52E-05	5.19E-05	3.42E-05
	Avg. a0	Avg. a1	Avg. a2	Avg. a3	Avg. a4	Avg. a5
	2.32E-05	0.34386	0.127008	0.0337753	0.094211	0.0649336

Table 8. We report forecasting regressions of integrated variance (estimated using optimally-defined flat-top cubic kernels as in Bandi and Russell, 2005) on five lags of cubic kernel estimates. The regressor's autocovariances are chosen using 3 methods described in the main text. We use Spiders mid-quotes on the NYSE. The sample period is 2002/1 - 2006/3. In all cases, 1,000 observations are employed to estimate the model parameters (a0 through a5) and forecast. The table reports means and standard deviations for both the regressand (true) and the volatility forecasts. It also reports the average number of autocovariances selected according to each rule.

LHS = FlatTukey RHS = FlatTukey 5 lags
2002/1-2006/3

(1)	Avg. q	MSE	Mean(Vtrue)	Mean(Vfore)	Std(Vtrue)	Std(Vfore)
	13.7749	**9.10E-10**	4.81E-05	5.49E-05	5.10E-05	3.67E-05
	Avg. a0	Avg. a1	Avg. a2	Avg. a3	Avg. a4	Avg. a5
	1.94E-05	0.372265	0.150015	0.0671072	0.091612	0.0789061
(2)	Avg. q	MSE	Mean(Vtrue)	Mean(Vfore)	Std(Vtrue)	Std(Vfore)
	19.8585	**1.0753E-09**	4.81E-05	5.09E-05	5.10E-05	3.24E-05
	Avg. a0	Avg. a1	Avg. a2	Avg. a3	Avg. a4	Avg. a5
	2.34E-05	0.308512	0.116799	0.0170259	0.087551	0.0556742
(3)	Avg. q	MSE	Mean(Vtrue)	Mean(Vfore)	Std(Vtrue)	Std(Vfore)
	7.94207	**1.0908E-09**	4.81E-05	5.45E-05	5.10E-05	3.38E-05
	Avg. a0	Avg. a1	Avg. a2	Avg. a3	Avg. a4	Avg. a5
	2.27E-05	0.341694	0.131405	0.0328646	0.089792	0.0768419

Table 9. We report forecasting regressions of integrated variance (estimated using optimally-sampled flat-top modified Tukey-Hanning kernels as in Bandi and Russell, 2005) on five lags of flat-top modified Tukey-Hanning kernel estimates. The regressor's autocovariances are chosen using 3 methods described in the main text. We employ Spiders mid-quotes on the NYSE. The sample period is 2002/1 - 2006/3. In all cases, 1,000 observations are employed to estimate the model parameters (a0 through a5) and forecast. The table reports means and standard deviations for both the regressand (true) and the volatility forecasts. It also reports the average number of autocovariances selected according to each rule.