Capturing Volatility from Large Price Moves: Generalized Range Theory and Applications

Dobrislav Dobrev[†]

Job Market Draft: November 25, 2006
This Draft: March 25, 2007

Abstract

I introduce a novel high-frequency volatility estimator based on large price moves, which constitutes a generalization of the range. Just like the standard range maximizes a single price difference on the observed price path, the generalized range maximizes the sum of multiple price differences. It provides a new link between recent advances in high-frequency volatility estimation and the long-established efficiency of the range. I develop an asymptotic theory for the generalized range in a jump-diffusion setting, originating a family of strongly consistent diffusive volatility estimators that are robust to jumps. This theory tackles maximization on a time grid not previously studied in the volatility literature and uncovers valuable distributional properties of price peaks and troughs.

On simulated data, the generalized range behaves in accordance with the derived theory and compares favorably to other known estimators of diffusive volatility that are robust to jumps. On real data, the generalized range is largely robust to microstructure noise when calculated on bid-ask quotes and proves valuable for intraday jump detection and short-term forecasting of stock return volatility. In a model-free environment, the capability of the generalized range to identify large zig-zag price moves appears to be directly applicable to relative value arbitrage strategies.

[†]Department of Finance, Kellogg School of Management, Northwestern University, Evanston, IL 60208, Tel.: +1 224 392 0782; Email: d-dobrev@kellogg.northwestern.edu

I would like to thank Torben Andersen, Ravi Jagannathan, Robert McDonald, and Ernst Schaumburg for many helpful discussions and suggestions. I am grateful for valuable remarks also to Luca Benzoni, Oleg Bondarenko, Robert Korajczyk, Jacobs Sagi, Costis Skiadas, Arne Staal, Rakesh Vohra as well as seminar participants at Barclays Global Investors, Board of Governors of the Federal Reserve System, Carnegie Mellon University, Federal Reserve Bank of Boston, Federal Reserve Bank of Chicago, Lehman Brothers, Northwestern University, University of Chicago, University of Michigan, and 2006 CIREQ Conference on Realized Volatility. All errors and omissions are my own.

1 Introduction

The importance of volatility for asset pricing, portfolio choice, and risk management coupled with the growing availability of high-frequency quotes and transaction prices has spurred substantial developments in volatility measurement, modeling, and forecasting beyond the traditional models based on daily data such as GARCH. The primary workhorse for this progress has been the realized volatility (RV) measure of the total daily price variance defined as the sum of squared consecutive intraday returns and valid under most general assumptions for the log-price process. Its theoretical and empirical merits have been scrutinized by Andersen and Bollerslev (1998), Andersen, Bollerslev, Diebold and Labys (2001, 2003), Barndorff-Nielsen and Shephard (2002a,b), and Meddahi (2002) among others. The underlying intuition that volatility can be estimated arbitrarily well by increasing the sampling frequency goes back to Merton (1980). In reality, however, because of market microstructure effects RV is reliable up to about a five minute sampling frequency. Important recent extensions of RV successfully deal with improving its efficiency in the presence of microstructure noise by means of optimal sampling, subsample averaging, or kernel-based estimation (see Aït-Sahalia, Mykland and Zhang , 2005a,b; Zhang, Aït-Sahalia and Mykland, 2005; Barndorff-Nielsen, Hansen, Lunde and Shephard, 2005b, 2006; Bandi and Russel, 2006; Hansen and Lunde, 2006, and references therein).

While RV is the benchmark high-frequency measure of total price variance, there is an important difference between the variance of continuous price changes due to normal information flow and the variance generated by price discontinuities upon the arrival of surprising news or abnormal orders. Separate measurement of the more predictable continuous part and less predictable jump part of total price variance is of interest for applications such as option pricing. Notable recent extensions of RV to bipower and multi-power variations allow direct measurement of the diffusive part of total price variance and disentangling the contribution of jumps (Barndorff-Nielsen and Shephard, 2004, 2006; Huang and Tauchen, 2005). This has allowed the development of novel procedures for detection of individual jumps (Andersen, Bollerslev, Frederiksen and Nielsen, 2006; Andersen, Bollerslev and Dobrev, 2006) that uncover previously unmeasurable features of asset prices. Explicit differentiation of the jump and continuous sample path components leads to gains in terms of volatility forecast accuracy (Andersen, Bollerslev and Diebold, 2005; Andersen, Bollerslev and Huang, 2006).

As far as volatility estimators based on a single intraday price move are concerned, it has long been known (Parkinson, 1980) that the properly scaled daily range or squared range are more efficient than the daily absolute and squared return. This insight has

inspired numerous further studies such as Garman and Klass (1980), Kunitomo (1992), Gallant, Hsu and Tauchen (1999), Alizadeh, Brandt and Diebold (2002), and more recently Dijk and Martens (2005) and Christensen and Podolskij (2006a,b). A useful way to appreciate Parkinson's result is that when volatility has to be estimated from just a single price move but more observations are available, it is best to take the largest price move among all the observed. Intuitively, the largest price moves convey a stronger volatility signal but their properties are more intricate than those of average price moves.

This paper shows that the same intuition generalizes to modern volatility estimators based on multiple intraday price moves. Roughly speaking, provided that the number of price moves that should be taken to estimate volatility is smaller than the number of available observations, it is beneficial to use the corresponding largest non-overlapping price moves between peaks and troughs on the observed path. This is quite appealing for applications on real market data since, regardless of how many observations are available, the presence of microstructure noise typically imposes a serious restriction on the number of intraday price moves that can be exploited without substantial measurement distortion. From this perspective, resorting exclusively to the largest price moves in the course of the day suggests a possible new way to minimize the efficiency loss caused by microstructure noise.

In what follows, I formalize this intuitive idea by means of a novel high-frequency generalization of the range. In the same way as the standard range maximizes a single price move on the observed price path, the generalized range (GR) maximizes the sum of any given number of non-overlapping price moves. The main effort in this paper is to establish a formal link between GR and volatility in a theoretically sound setting. In particular, assuming that the logarithmic price follows a generic jump-diffusion (with finite number of jumps in each time interval), the obtained core theoretical result is that GR can be properly scaled to define novel strongly consistent and asymptotically unbiased estimators of the diffusive part of total price variance. The involved maximization on a time grid does not allow building a close parallel to the existing theory for RV. In order to derive the almost sure asymptotics of GR, I uncover and utilize distributional properties of price peaks and troughs (formally defined as Brownian h-extrema), which as a useful new tool for analysis makes a contribution of its own to the volatility literature.

As an estimator of the diffusive part of total price variance, GR delivers a new viable alternative to the bipower variation of Barndorff-Nielsen and Shephard (2004) and the closely related bipower range of Christensen and Podolskij (2006b). A realistic finite sample experiment suggests that GR compares favorably to both alternatives in accordance with the underlying intuition. Moreover, calculation of GR on bid-ask quotes leads

to a natural selection of the maximum number of price moves that can be used without distortion by microstructure noise. Improving the measurement of the diffusive volatility component is of particular help to recently developed jump detection procedures as in Andersen, Bollerslev, and Dobrev (2006) and volatility forecasting frameworks as in Andersen, Bollerslev, and Diebold (2005). An application of GR to short term stock volatility forecasting provides indirect evidence that GR performs reasonably well as an estimator of the diffusive part of total price variance not only on simulations but also on real market data.

Last but not least, the unique features of GR are valuable in finance applications even without making particular assumptions on the dynamics of the price process or the sampling scheme. In particular, the capability of GR to capture large zig-zag price moves appears to be useful for relative value arbitrage strategies such as pairs-trading considered in Gatev, Goetzmann, and Rouwenhorst (2006). In accordance with the intuition behind its very definition, the generalized range can simply be used to identify profitable stock pairs whose price difference follows a convergence/divergence pattern of larger amplitudes.

The paper proceeds as follows. Section 2 formally defines the generalized range and provides some additional interpretation. Section 3 builds an asymptotic theory for the generalized range and shows that it can be properly scaled to define novel strongly consistent and asymptotically unbiased estimators of the diffusive part of total price variance in a jump-diffusion setting. Section 4 documents that the generalized range compares favorably to its bipower variation alternatives on simulated data. Section 5 carries out an application of the generalized range to volatility forecasting on stock data. Section 6 illustrates in the context of relative value arbitrage strategies that GR appears to be useful also in a model-free environment. Section 7 concludes.

2 Definition and Interpretation of the Generalized Range

Given a path $\mathbf{p}$ of logarithmic asset prices (cumulative asset returns) on an interval $[0, T]$, I define the generalized range $\mathbf{GR}_k$ of order k as the sum of the magnitudes of the largest k non-overlapping price moves (ups and downs) along the path. Formally,

$$\mathbf{GR}_k(\mathbf{p}_{[0,T]}) = \max_{0 \le t_1 \le \cdots \le t_{2k} \le T} \sum_{i=1}^{k} \left| \mathbf{p}(t_{2i}) - \mathbf{p}(t_{2i-1}) \right| . \tag{1}$$

Figure 1 illustrates the definition of $\mathbf{GR}$ for order $k = 4$, i.e. four price moves. $\mathbf{GR}$ has an appealing intuitive interpretation as the maximum trading gain or loss that can

be attained by executing k consecutive and non-overlapping intraday trades (buy/sell to open a position and then sell/buy to close it) given the price evolution. In particular, from a risk measurement perspective, **GR** captures the outcome of the worst possible timing of k trades (buying at peaks and selling at troughs), which can be interpreted as a natural generalization of the maximum drawdown measure of risk used by practitioners[1].

GR of order $k = 1$ reduces to the standard range (high minus low), which has been extensively studied as a volatility estimator starting with Parkinson (1980). **GR** of order $k > 1$ resembles a power variation measure with k terms, however, it has two distinct new features: first, **GR** involves maximization on the time grid, and second, **GR** allows but does not require the time intervals of its k terms to be adjacent[2].

These key features render **GR** a natural high-frequency generalization of the standard daily range. It is substantially different from the range-based extensions of realized volatility concurrently pursued by Dijk and Martens (2005) and Christensen and Podolskij (2006a,b) who take sums of powers or adjacent products of multiple intraday ranges on a fixed time grid. Such extensions exploit properties of the standard range within the established framework of realized volatility. Similarly, one can exploit properties of **GR** within the same framework by substituting the standard range with **GR** in each intraday interval on the grid. However, first one needs to establish **GR** as a volatility estimator, which is the main subject of this paper.

The calculation of **GR** is well-suited for standard optimization packages as the arising maximization problem on a discrete time grid of price observations has the interpretation of finding a maximum flow through a network with gains. More efficient algorithms exploiting ad-hoc properties of the particular maximization problem at hand have been developed by Dobrev (1998).

3 Theoretical Foundations

In what follows, the logarithmic asset price process is assumed to evolve according to a generic jump-diffusion with jumps of finite activity. The continuous component is meant to capture the price changes due to normal market activity (information and transaction

[1] A quick overview of risk-adjusted measures of performance based on the maximum drawdown can be found in Vecer (2006) and Magdon-Ismail and Atiya (2004).

[2] It is easy to see that **GR** defines a norm on the space of log-price process starting at the origin. Also, the generalized range of any power $q \geq 1$ can be defined as $\mathbf{GR}_k^{(q)}(\mathbf{p}_{[0,T]}) = \max_{0 \leq t_1 \leq \cdots \leq t_{2k} \leq T} \sum_{i=1}^{k} |\mathbf{p}(t_{2i}) - \mathbf{p}(t_{2i-1})|^q$. This paper studies **GR** of power $q = 1$ since it provides the most direct high-frequency analog to the standard daily range, while $q > 1$ poses additional issues arising from convexity.

flow), while jumps reflect the arrival of surprising news or abnormally large orders. This setting is theoretically sound since as in Back (1991) it naturally leads to the conclusion that the price process is a special semi-martingale, ruling out arbitrage.

It is worth noting also that a number of recent studies have confirmed that, unlike pure diffusions, jump-diffusion settings of this kind are largely consistent with various financial return series (see Andersen, Benzoni and Lund, 2002; Carr and Wu, 2003; Eraker 2004; Eraker, Johannes, and Polson, 2003; Maheu and McCurdy, 2004 and Pan, 2002 among others). Moreover, recently developed procedures for testing the general adequacy of jump-diffusion models for return distributions as in Andersen, Bollerslev, and Dobrev (2006) and Andersen, Bollerslev, Frederiksen and Nielsen (2006) have not been able to formally rule out settings of this kind, while casting serious doubt on pure diffusion models without jumps.

3.1 Setup and Preview of Main Theory Contributions

In view of the above remarks, without much loss of generality, the log-price process $\mathbf{p}(t)$ will be assumed to follow a standard jump-diffusion governed by the stochastic differential equation

$$d\mathbf{p}(t) = \mu(t)dt + \sigma(t)dW(t) + \kappa(t)dq(t),\ t \in [0,T], \tag{2}$$

where the volatility process $\sigma(t)$ is strictly positive (bounded away from zero) and càdlàg, $W(t)$ is a Wiener process, $q(t)$ is a finite-activity counting process with $dq(t) = 1$ whenever there is a jump of size $\kappa(t)$ at time t, and $dq(t) = 0$ otherwise. The finite jump intensity will be denoted as $\lambda(t)$, allowing for possible time variation. The drift $\mu(t)$ is a bounded variation process and it is easy to show that ignoring it will not affect the asymptotic results obtained below.[3]

The total variability of the cumulative return process $\mathbf{p}$ on the interval $[0,T]$ is given by the latent quadratic variation $QV_{[0,T]}$. In this setting it can be expressed as the sum of distinct pure diffusive and jump components:

$$QV_{[0,T]} = \int_0^T \sigma^2(s)ds + \sum_{0<s\leq T} \kappa(s)^2 = IV_{[0,T]} + JV_{[0,T]}. \tag{3}$$

In particular, the diffusive component is defined by the integrated variance $IV_{[0,T]} = \int_0^T \sigma^2(s)ds$, while the jump contribution is simply the sum of squared jumps $JV_{[0,T]} = \sum_{0<s\leq T} \kappa(s)^2$.

[3]This point is formalized in the appendix as part of the proof of proposition 7 below.

Much of the recent work in the area of volatility measurement, modeling, and forecasting deals with improving the inference of these two volatility components based on high-frequency data. First, as a benchmark measure of total volatility has been established realized volatility (RV) defined as the sum of squared consecutive intraday returns, with a few notable recent extensions that successfully deal with microstructure noise (see Andersen and Bollerslev, 1998; Andersen, Bollerslev, Diebold and Labys, 2001, 2003; Barndorff-Nielsen and Shephard, 2002a, b; and Meddahi, 2002; Aït-Sahalia, Mykland, and Zhang, 2005a,b; Barndorff-Nielsen, Hansen, Lunde and Shephard, 2005b; and Hansen and Lunde, 2006, among others). Second, the benchmark tool for measuring the pure diffusive volatility in the presence of jumps has been the bipower variation (BV) defined as the sum of adjacent products of absolute intraday returns (see Barndorff-Nielsen and Shephard, 2004, 2006; Andersen, Bollerslev, and Diebold, 2005; Huang and Tauchen, 2005), with a very recent extension to an analogous bipower range measure, which takes the sum of products of adjacent ranges rather than absolute returns (see Christensen and Podolskij, 2006b).

In this light, the main contribution of this paper to the current volatility literature is that **GR** can be naturally scaled to define a novel measure of the pure diffusive volatility component in the adopted standard jump-diffusion setting. It compares favorably to both bipower variation alternatives, especially in the presence of microstructure noise when the largest price moves exclusively picked by **GR** tend to remain relatively more robust to noise than those picked by the bipower variation measures.

It is worth noting also that the involved maximization on the time grid does not allow building the asymptotic theory for **GR** in close parallel to the existing theory for RV. In fact, one of the new auxiliary tools that I exploit in the main derivations that follow, namely the Brownian h-extrema (a formal definition is given below), makes a contribution of its own to the volatility literature. It is further explored in concurrent work in progress developing a novel duration-based approach to volatility estimation based on the passage times between price peaks and troughs that differ by more than a certain threshold.

3.2 Asymptotic Theory for GR In The Absence of Jumps and Drift

In what follows, **GR** calculated on n observations of $\mathbf{p}$ in the interval $[0, T]$ will be denoted as $\mathbf{GR}_k^n(\mathbf{p}_{[0,T]})$. The notation is purposely simplified not to indicate the sample mesh size (the maximum distance between adjacent observations), since such extra requirements will be explicitly stated whenever needed. The infeasible value of **GR** calculated on the whole unobserved path of $\mathbf{p}$ on $[0, T]$ (rather than a finite number of observations) will be

denoted as $\mathbf{GR}_k^{\ominus}(\mathbf{p}_{[0,T]})$[4]. As in the case of the standard daily range, the fewer are the available observations, the larger is the downward bias of the feasible value $\mathbf{GR}_k^n(\mathbf{p}_{[0,T]})$ relative to the corresponding infeasible value $\mathbf{GR}_k^{\ominus}(\mathbf{p}_{[0,T]})$. Nonetheless, the superscripts n and $\ominus$ will be suppressed in statements for $\mathbf{GR}$ that are valid in both cases. Likewise, whenever there is no ambiguity, the process $\mathbf{p}$ and the time interval $[0,T]$ may also be suppressed to simplify notation.

Using this notation, I summarize pathwise properties of $\mathbf{GR}$ in the absence of jumps and use them to establish the asymptotic behavior of $\mathbf{GR}$ as its order k gets large, i.e. when more and more non-overlapping price moves of largest possible magnitude are taken within a given interval of time.

3.2.1 Pathwise Properties of GR on a Brownian Diffusion

In the absence of jumps, the logarithmic price process in (2) reduces to a Brownian diffusion and the total variability of the cumulative return process $\mathbf{p}$ is entirely due to the pure diffusive volatility given by the integrated variance $IV_{[0,T]} = \int_0^T \sigma^2(s)ds$. What makes $\mathbf{GR}$ tractable as a measure of the integrated variance $IV_{[0,T]}$ in this setting is the invariance of the maximization operator with respect to continuous non-decreasing time change and the fact that $\mathbf{p}$ can be represented as a time-changed Brownian motion $\mathbf{b} \circ \boldsymbol{\tau}$ a.s. by the Dambis-Dubins-Schwarz theorem (Dambis, 1965, and Dubbins and Schwartz, 1965). For a Brownian diffusion the time change $\boldsymbol{\tau}$ is given exactly by the continuous non-decreasing integrated variance process of $\mathbf{p}$ defined as:

$$\boldsymbol{\tau}(t) = \int_0^t \sigma^2(s)ds \equiv IV_{[0,t]}, t \in [0,T]$$

The economic interpretation of $\boldsymbol{\tau}$ is a time clock that measures financial time, business time, or transaction time, as motivated by Clark (1973) and more recently Ane and Geman (2000).

Thanks to monotonicity and continuity of the time change τ, the definition of $\mathbf{GR}$ implies that

$$\mathbf{GR}_k(\mathbf{p}_{[0,T]}) = \mathbf{GR}_k(\mathbf{b}_{[0,\boldsymbol{\tau}(T)]})$$

In words, $\mathbf{GR}$ for the log-price process $\mathbf{p}$ on the physical time interval $[0,T]$ coincides with $\mathbf{GR}$ for the Brownian motion $\mathbf{b}$ on the corresponding financial time interval $[0,\boldsymbol{\tau}(T)]$,

[4]Note that in the absence of jumps the paths of $\mathbf{p}$ are continuous a.s. and hence the infeasible value $\mathbf{GR}_k^{\ominus}(\mathbf{p}_{[0,T]})$ is well-defined and finite a.s.

whose end point $\boldsymbol{\tau}(T)$ is exactly the integrated variance $IV_{[0,T]}$ that we want to measure. By using the scaling property of Brownian motion it is straightforward to obtain (see appendix) also that

$$\mathbf{GR}_k(\mathbf{b}_{[0,\boldsymbol{\tau}(T)]}) = \sqrt{IV_{[0,T]}} \times \mathbf{GR}_k(\mathbf{B}_{[0,1]}),$$

where $\mathbf{B}_{[0,1]}$ stands for some standard Brownian path on the unit interval $[0,1]$. Combining the above two equalities leads to the following main pathwise property of **GR**:

P1 - Variance Separability: *On each path of the log-price process* $\mathbf{p}$ *in* $[0,T]$*,* **GR** *scales by a factor of* $\sqrt{IV_{[0,T]}}$ *the value of* **GR** *on a corresponding (via time-change and scaling) standard Brownian path* $\mathbf{B}$ *in the interval* $[0,1]$*:*

$$\mathbf{GR}_k(\mathbf{p}_{[0,T]}) = \sqrt{IV_{[0,T]}} \times \mathbf{GR}_k(\mathbf{B}_{[0,1]})$$

Thus, knowledge of **GR** on the standard Brownian paths in $[0,1]$ would suffice for disentangling $\sqrt{IV_{[0,T]}}$ from the value of **GR** on the paths of the process $\mathbf{p}$ in $[0,T]$. In essence, it is sufficient to focus on the derivation of the almost sure asymptotics of **GR** on the Brownian path space in $[0,1]$ and then use the above key property to define a family of strongly consistent diffusive volatility estimators based on **GR**. Before doing so, though, it is useful to state also a few other pathwise properties of **GR** that will play a role in the analysis.

P2 - Lower Range-Based Bound: $\mathbf{GR}_k \geq \sum_{i=1}^{k} \max_{s_{k-1} \leq u < v \leq s_k} |\mathbf{p}(u) - \mathbf{p}(v)|$ *for any grid of* $k+1$ *times* $0 = s_0 < s_1 < \cdots < s_k = T$*, i.e.* $\mathbf{GR}_k$ *cannot be lower than the sum of the ranges of* $\mathbf{p}$ *calculated within any chosen* k *consecutive intervals in* $[0,T]$*.*

Clearly, this is a direct implication from the definition of **GR**. More generally, let $0 \leq t_1^* \leq t_2^* \leq \cdots \leq t_{2k}^* \leq T$ be optimal times for $\mathbf{GR}_k(\mathbf{p}_{[0,T]})$, i.e. $\sum_{i=1}^{k} \left|\mathbf{p}(t_{2i}^*) - \mathbf{p}(t_{2i-1}^*)\right| \equiv \mathbf{GR}_k(\mathbf{p}_{[0,T]})$. Then the involved maximization on the time grid implies also the following important characteristics of **GR**:

P3 - Extremality: *The optimal times* $t_1^* \leq t_2^* \leq \cdots \leq t_{2k}^*$ *represent points of alternating local maxima and minima for the observations of* $\mathbf{p}$ *in* $[0,T]$*. Moreover, for any two times* u *and* v *of a local maximum and minimum that can be augmented to the optimal times* $t_1^* \leq t_2^* \leq \cdots \leq t_{2k}^*$ *without destroying the alternation of the corresponding extrema, it must be the case that* $|\mathbf{p}(u) - \mathbf{p}(v)| \leq \min_{i=1,\ldots,k} \left|\mathbf{p}(t_{2i}^*) - \mathbf{p}(t_{2i-1}^*)\right|$.

P4 - Monotonicity: $\mathbf{GR}_{k+1} \geq \mathbf{GR}_k$ *with strict inequality whenever the times* $t_1^* \leq t_2^* \leq \cdots \leq t_{2k}^*$ *do not exhaust all local extrema among the observations of* $\mathbf{p}$ *in* $[0,T]$.

P5 - Non-Increasing Differences: $\mathbf{GR}_{k+1} - \mathbf{GR}_k \leq \mathbf{GR}_k - \mathbf{GR}_{k-1}$, *i.e. the first order differences of* $\mathbf{GR}_k$ *with respect to* k *are non-increasing.*

P6 - Dense Coverage: *The shortest of the* k *non-overlapping time intervals* $t_2^* - t_1^*,\ t_4^* - t_3^*,\ \cdots,\ t_{2k}^* - t_{2k-1}^*$ *cannot exceed* T/k. *Moreover, a.s. the largest of these intervals also shrinks to zero as* $k \longrightarrow \infty$, *provided that the full path of* $\mathbf{p}$ *in* $[0,T]$ is *observed (or the mesh size of the finite sample shrinks to zero faster enough than* $1/k$*).*

The last property follows by the fact that a.s. each path of $\mathbf{p}$ is nowhere monotone on the finite interval $[0,T]$, while it is subject to a uniform modulus of continuity as detailed in the next section. Thus, roughly speaking, when k gets large the growth of $\mathbf{GR}$ will be driven by the largest remaining price moves between local maxima and minima that are getting closer and closer to each other.

The obtained key properties of $\mathbf{GR}$ set the stage for establishing its almost sure asymptotics on a standard Brownian motion.

3.2.2 Almost Sure Asymptotics of GR on a Standard Brownian Motion

Thanks to variance separability (property P1), it suffices to establish the exact almost sure asymptotics of $\mathbf{GR}_k$ as $k \longrightarrow \infty$ on the space Ω of the unit interval paths of a standard Brownian motion $\mathbf{B}$. To this end, it is convenient to use the a.s. representation of Ω as all continuous functions $\omega : [0,\infty] \longrightarrow \mathbb{R}$ restricted to the unit interval and equipped with the Wiener measure. What needs to be shown is that for almost every Brownian path $\omega \in \Omega$ the asymptotic behavior of the sequence $\{\mathbf{GR}_k(\omega)\}_{k=1}^{\infty}$ is the same.

It is convenient to focus first on the infeasible $\mathbf{GR}_k^{\ominus}(\omega)$ calculated on a fully observed Brownian path ω on the unit interval and then extend the obtained results to the feasible analog $\mathbf{GR}_k^n(\omega)$ calculated on n observations of ω subject to a suitable restriction on the sampling scheme.

Extremality of $\mathbf{GR}$ (property P3) implies that the limiting properties of $\left\{\mathbf{GR}_k^{\ominus}(\omega)\right\}_{k=1}^{\infty}$ are determined by the set of local extrema on each Brownian path $\omega \in \Omega$. It is a well known fact that the Brownian local extrema in $[0,1]$ are a.s. a countable and everywhere dense subset of $[0,1]$. Besides, all local extrema are strict and never take the same value (see, for example, Karatzas and Shreve, 1991, or Revuz and Yor, 1999). In particular, this means that a.s. there is no discrepancy between the cardinality of the set of all Brownian local extrema on the path ω and the cardinality of the set of local extrema picked by

$\mathbf{GR}_k^{\ominus}(\omega)$ as $k \longrightarrow \infty$. Moreover, strict monotonicity (property P4) and dense coverage (property P6) in this case imply also that a.s. $\mathbf{GR}_{k+1}^{\ominus} > \mathbf{GR}_k$ for any k and the set of local extrema picked by $\mathbf{GR}_k^{\ominus}(\omega)$ as $k \longrightarrow \infty$ should in fact coincide with the entire set of local extrema on the path $\omega \in \Omega$. Nonetheless, not enough is known about the entire set of Brownian local extrema to deduce the exact asymptotics of $\mathbf{GR}_k^{\ominus}(\omega)$ by an argument of this kind.[5]

Instead, I resort to the main insights from the previous section. In particular, extremality (property P3) implies that if

$$\mathbf{GR}_k^{\ominus} = \sum_{i=1}^{k} \left| \mathbf{B}(t_{2i}^*) - \mathbf{B}(t_{2i-1}^*) \right|$$

then the optimal times $t_1^* \leq t_2^* \leq \cdots \leq t_{2k}^*$ taken without repetitions should form a sequence of $k+q_k+1$ distinct times $s_1^* < s_2^* < \cdots < s_{k+q_k+1}^*$ such that $\mathbf{B}_{s_1^*}$, $\mathbf{B}_{s_2^*}$, $\cdots$, $\mathbf{B}_{s_{k+q_k+1}^*}$ are alternating Brownian local maxima and minima, where q_k is some integer between 0 and $k-1$. Moreover, each s_i^* must be a point of global maximum or minimum on the interval $[s_{i-1}^*, s_{i+1}^*]$. Hence, if the threshold g_k is defined as the smallest among the $k+q_k$ absolute differences between these alternating local maxima and minima, i.e.

$$g_k = \min_{i=1\ ,\ \ldots\ ,\ k+q_k} \left| \mathbf{B}_{s_{i+1}^*} - \mathbf{B}_{s_i^*} \right|,$$

then $\mathbf{B}_{s_1^*}$, $\mathbf{B}_{s_2^*}$, $\cdots$, $\mathbf{B}_{s_{k+q_k+1}^*}$ must in fact be a sequence of Brownian h-extrema for threshold level $h = g_k$. A point s_h is said to be an h-minimum (maximum) of a function if it is contained in an interval on which s_h is the global minimum (maximum) and the value taken by the function at both end points of the interval is above (below) this global minimum (maximum) at least by a threshold h.[6] The concept of h-extrema is illustrated by figure 2, while the following claim formalizes the tight link between Brownian h-extrema and $\mathbf{GR}_k^{\ominus}$:

Claim 1 (Link between Brownian h-extrema and $\mathbf{GR}_k^{\ominus}$) $\mathbf{GR}_k^{\ominus}$ *is the sum of the largest k among $k+q_k$ absolute differences between adjacent Brownian g_k-extrema for an*

[5] The lack of sufficient knowledge about the entire set of Brownian local extrema is largely because they are not well suited for applying standard tools for analysis such as the reflection principle. A very recent study by Tsirelson (2006) sheds a bit more light on this by demonstrating that a.s. the times of all Brownian local extrema on $[0,1]$ coincide with a certain reshuffling of a randomly drawn infinite sequence of independent and uniformly distributed numbers on $[0,1]$.

[6] Formally, s_h is said to be an h-minimum of a function $f(\cdot)$ if there exist points s_h^- and s_h^+ such that $s_h^- < s_h < s_h^+$, $f(s_h^-) \geq f(s_h) + h$, $f(s_h^+) \geq f(s_h) + h$, and $f(s_h) = \min_{[s_h^-, s_h^+]} f$. Similarly, s_h is an h-maximum of f if s_h is an h-minimum of $-f$.

appropriate threshold level g_k that ensures exactly $k+q_k+1$ Brownian g_k-extrema in the interval $[0,1]$, where q_k is some integer between 0 and k.

This characterization of $\mathbf{GR}_k^{\ominus}$ is very helpful, since the distributional properties of consecutive Brownian local extrema above any fixed threshold have been fully described by Neveu and Pitman (1989). More precisely, for any fixed threshold $h>0$, the Brownian h-extrema $\left\{\mathbf{B}(s_{j,[h]})\right\}_{j=0}^{\infty}$ and their times $\left\{s_{j,[h]}\right\}_{j=0}^{\infty}$ have the following nice renewal property:

1. $\left|\mathbf{B}(s_{j,[h]})-\mathbf{B}(s_{j-1,[h]})\right|-h,\ \ j=1,2,...$ are independent and identically distributed exponentially, with parameter h;

2. $s_{j,[h]}-s_{j-1,[h]},\ \ j=1,2,...$ are independent and identically distributed, with Laplace transform $\psi(x)=1/\cosh(h\sqrt{2x})$.

In particular, the mean and variance of the price moves between Brownian h-extrema and their durations are:

$$\begin{aligned}\mathbb{E}\left[\left|\mathbf{B}(s_{j,[h]})-\mathbf{B}(s_{j-1,[h]})\right|\right] &= 2h,\ \ \mathbb{E}\left[s_{j,[h]}-s_{j-1,[h]}\right]=h^2,\ \ j=1,2,...\\ \mathbb{V}\left[\left|\mathbf{B}(s_{j,[h]})-\mathbf{B}(s_{j-1,[h]})\right|\right] &= h^2,\ \mathbb{V}\left[s_{j,[h]}-s_{j-1,[h]}\right]=h^4,\ \ j=1,2,... \qquad (4)\end{aligned}$$

Note that in claim 1 above for the interval $[0,1]$ to contain $k+q_k+1$ Brownian g_k-extrema, the threshold g_k must go down if q_k goes up. Given the renewal property of Brownian h-extrema, this essentially implies that the maximization in $\mathbf{GR}_k^{\ominus}$ involves an implicit trade-off between a larger number q_k and a smaller mean g_k when taking the k largest among $k+q_k$ i.i.d. exponential random variables with mean g_k (taking into account also the shift of all of them by g_k).

In essence, all this implies that the limiting behavior of $\mathbf{GR}_k^{\ominus}$ will depend on the limiting behavior of g_k versus q_k and the arising exponential order statistics. Lemma 9 in the appendix exploits the durations between Brownian h-extrema to place an important restriction on the joint asymptotics of g_k and $k+q_k$. Then based on this restriction, lemma 10 and its corollary 11 provide a limsup result for $\mathbf{GR}_k^{\ominus}$ by exploiting known characteristics of exponential order statistics. Finally, lemma 12 provides an identical liminf result for $\mathbf{GR}_k^{\ominus}$, which leads to the following core proposition:

Proposition 2 (Almost Sure Asymptotics of $\mathbf{GR}_k^{\ominus}$) $\frac{1}{2}k^{-\frac{1}{2}}\mathbf{GR}_k^{\ominus}\xrightarrow{a.s.}1$

Corollary 3 *By L_2-boundedness* $\frac{1}{2}k^{-\frac{1}{2}}\mathbb{E}\left|\mathbf{GR}_k^{\ominus}\right|\longrightarrow 1$, *so* $\left(\mathbb{E}\left|\mathbf{GR}_k^{\ominus}\right|\right)^{-1}\mathbf{GR}_k^{\ominus}\xrightarrow{a.s.}1$

The next step is to extend the derived asymptotic results for the infeasible $\mathbf{GR}_k^{\ominus}$ to similar results for the feasible $\mathbf{GR}_k^{n}$ calculated on n observations instead of the whole path.

This can be achieved by imposing double large sample asymptotics $k \longrightarrow \infty$, $n \longrightarrow \infty$ such that the sample mesh size Δ_n (i.e. the maximum distance between adjacent observations) shrinks to zero faster enough than $1/k$ to make $\mathbf{GR}_k^n$ get closer and closer to $\mathbf{GR}_k^{\ominus}$ a.s. in the limit[7]. One way to formalize this idea is to exploit Levy's uniform modulus of continuity for standard Brownian motion (see, for example, Karatzas and Shreve, 1991, or Revuz and Yor, 1999):

$$\lim_{\varepsilon \downarrow 0} \sup_{\substack{0 \leq t \leq 1 \\ 0 < s < \varepsilon}} \frac{|\mathbf{B}_{t+s} - \mathbf{B}_t|}{(2s\,|\log s|)^{1/2}} = 1 \quad a.s.$$

In essence, Levy's uniform modulus of continuity implies an a.s. upper bound on the discrepancy between $\mathbf{GR}_k^n$ and $\mathbf{GR}_k^{\ominus}$ given a suitable restriction on the mesh size Δ_n and leads to the following large sample analog to proposition 2 (leaving again all details for the appendix):

Proposition 4 (Almost Sure Asymptotics of $\mathbf{GR}_k^n$) $\frac{1}{2}k^{-\frac{1}{2}}\mathbf{GR}_k^n \xrightarrow{a.s.} 1$ *provided that the mesh size* Δ_n *satisfies* $k\Delta_n\,|\log \Delta_n| \longrightarrow 0$ *as* $k \longrightarrow \infty$, $n \longrightarrow \infty$.

Corollary 5 *By* L_2*-boundedness* $\frac{1}{2}k^{-\frac{1}{2}}\mathbb{E}\,|\mathbf{GR}_k^n| \longrightarrow 1$, *so* $(\mathbb{E}\,|\mathbf{GR}_k^n|)^{-1}\,\mathbf{GR}_k^n \xrightarrow{a.s.} 1$ *provided that the mesh size* Δ_n *satisfies* $k\Delta_n\,|\log \Delta_n| \longrightarrow 0$ *as* $k \longrightarrow \infty$, $n \longrightarrow \infty$.

The obtained asymptotic results for **GR** on a standard Brownian motion along with the key variance separability property of **GR** are exploited in the next section to define novel diffusive volatility estimators based on **GR**.

3.3 Diffusive Volatility Estimators Based on GR

Going back to the general jump-diffusion setting of sde (2) above, it turns out that the obtained results for **GR** on a Brownian diffusion naturally lead to strongly consistent and asymptotically unbiased estimators of the diffusive volatility in the potential presence of finite number of jumps (and bounded variation drift). The argument is as follows.

First, when there are no jumps, variance separability (property P1) implies that on each path of the log-price process $\mathbf{p}$ in $[0, T]$, **GR** scales by a factor of $\sqrt{IV_{[0,T]}}$ the value of **GR** on a corresponding (via time-change and scaling) standard Brownian path $\mathbf{B}$ in

[7] Without loss of generality, it is also convenient to assume that no previously sampled observations are thrown out as n gets large, i.e. only new observations are added. This ensures that for each $k \geq 1$, $\mathbf{GR}_k^n \nearrow \mathbf{GR}_k^{\ominus}$ as $n \longrightarrow \infty$, i.e. the larger the sample size becomes, the closer the finite sample value of **GR** gets from below to its true value calculated on the whole path.

the interval $[0,1]$:

$$\mathbf{GR}_k(\mathbf{p}_{[0,T]}) = \sqrt{IV_{[0,T]}} \times \mathbf{GR}_k(\mathbf{B}_{[0,1]})$$

In this case, the established almost sure asymptotics of $\mathbf{GR}_k$ on the Brownian paths $\mathbf{B}$ in $[0,1]$ yields:

$$\mathbb{E}^{-1}\left[\mathbf{GR}_k(\mathbf{B}_{[0,1]})\right] \times \mathbf{GR}_k(\mathbf{B}_{[0,1]}) \xrightarrow{a.s.} 1$$

The scaling factor $e'_k \doteq \left(\mathbb{E}\left[\mathbf{GR}_k(\mathbf{B}_{[0,1]})\right]\right)^{-1} \sim 1/2k^{-1/2}$ is fully determined by the Brownian paths on $[0,1]$. Thus,

$$e'_k \times \mathbf{GR}_k(\mathbf{p}_{[0,T]}) \xrightarrow{a.s.} \sqrt{IV_{[0,T]}},$$

which means that in the absence of jumps $e'_k \times \mathbf{GR}_k(\mathbf{p}_{[0,T]})$ defines a strongly consistent estimator of $\sqrt{IV_{[0,T]}}$.

Second, the presence of a finite number of jumps can cause only a finite distortion of the corresponding value of $\mathbf{GR}_k(\mathbf{p}_{[0,T]})$ without jumps (for any k the distortion is in fact uniformly bounded by the sum of the magnitudes of all jumps, see details in the appendix in the proof of proposition 7 below). At the same time, as k gets large the scaling factor e'_k shrinks to zero at the rate of $k^{-1/2}$. Hence, the limiting value of the product $e'_k \times \mathbf{GR}_k(\mathbf{p}_{[0,T]})$ will not be affected by the jumps. Likewise, there will be no effect in the limit in case there is a bounded variation drift as well.

An analogous argument can be followed also after squaring both sides of the above pathwise equality or taking a logarithm on both sides, while using Slutsky's theorem to manipulate the a.s. limit accordingly. This gives rise to the following definitions of estimators of $\sqrt{IV_{[0,T]}}$, $IV_{[0,T]}$, and $\log\sqrt{IV_{[0,T]}}$ based on $\mathbf{GR}_k$:

Definition 6 (Estimators of $IV_{[0,T]}$, $\sqrt{IV_{[0,T]}}$, and $\log\sqrt{IV_{[0,T]}}$ based on GR) *Let the log-price process* $\mathbf{p}$ *on* $[0,T]$ *be governed by the generic jump-diffusion sde (2) and let* $IV_{[0,T]} = \int_0^T \sigma^2(s)ds$ *be the diffusive component of the volatility decomposition (3) in this setting. Then for all* $k = 1, 2, ...$ *the following estimators are well defined:*
1. An estimator of $\sqrt{IV_{[0,T]}}$ *is* $e'_k \times \mathbf{GR}_k(\mathbf{p}_{[0,T]})$*, where* $e'_k \doteq \mathbb{E}^{-1}\left[\mathbf{GR}_k(\mathbf{B}_{[0,1]})\right]$.
2. An estimator of $IV_{[0,T]}$ *is* $e''_k \times \left(\mathbf{GR}_k(\mathbf{p}_{[0,T]})\right)^2$*, where* $e''_k \doteq \mathbb{E}^{-1}[\left(\mathbf{GR}_k(\mathbf{B}_{[0,1]})\right)^2]$.
3. An estimator of $\log\sqrt{IV_{[0,T]}}$ *is* $e'''_k + \log \mathbf{GR}_k(\mathbf{p}_{[0,T]})$*, where* $e'''_k \doteq -\mathbb{E}\left[\log \mathbf{GR}_k(\mathbf{B}_{[0,1]})\right]$.
In particular, the scaling factors e'_k*,* e''_k*, and* e'''_k *are fully determined by the standard Brownian paths on* $[0,1]$ *and can be obtained with arbitrary precision by Monte Carlo simulation for any finite sample size* n*.*

The next two propositions summarize the main properties of these estimators**.**

Proposition 7 (Strong consistency) *The* **GR***-based diffusive volatility estimators defined above are strongly consistent as* $k \longrightarrow \infty$*, where in the large sample case the sample size* n *grows in accordance with the sampling condition of proposition 2.*

Proposition 8 (Asymptotic unbiasedness) *The* **GR***-based diffusive volatility estimators defined above are asymptotically unbiased as* $k \longrightarrow \infty$*, where in the large sample case the sample size* n *grows in accordance with the sampling condition of proposition 2.*

The proof detailed in the appendix shows that the presence of jumps (and possibly drift) is not the only source of asymptotically negligible bias. The other main driver of insignificant bias is the potential presence of the so called "leverage effect", commonly used to denote any type of dependency between the log-price and volatility innovations consistent with the empirically documented asymmetric relationship between them (see for example Bollerslev, Litvinova, and Tauchen, 2005, and references therein).

In any case, the other known diffusive volatility estimators in a jump-diffusion setting, namely the bipower variation of Barndorff-Nielsen and Shephard (2004, 2006) and its recent extension by Christensen and Podolskij (2006b), are also strongly consistent and asymptotically unbiased in the presence of jumps. Moreover, knowledge of the asymptotic variance of these estimators in the absence of jumps is not directly informative of their efficiency in the presence of jumps. Therefore, it is exclusively a matter of finite sample analysis to determine whether the diffusive volatility estimators based on **GR** can offer efficiency gains in a realistic jump-diffusion setting.[8] This is the main task of the following section.

4 Finite Sample Efficiency and Jump Signature Plots

Jumps in asset prices reflect the arrival of significant news or abnormal orders that have not been anticipated. As discussed in section 3, a number of parametric and recent non-parametric studies have concluded that models for asset prices without jumps cannot fit reasonably well neither stock returns nor out of the money option prices. Importantly, it

[8]Knowledge of the asymptotic variance of a diffusive volatility estimator under the null of no jumps is important for jump detection procedures along the lines of Barndorff-Nielsen and Shephard (2005a), Andersen, Bollerslev, and Diebold (2005), Huang and Tauchen (2005), Andersen, Bollerslev, Frederiksen, and Nielsen (2006). Finite sample efficiency of a diffusive volatility estimator in the presence of jumps is important for jump detection procedures along the lines of Andersen, Bollerslev, and Dobrev (2006).

has been found that option prices typically reveal larger jump risk than stock prices. In this regard, better identification of the realized jump volatility component is essential for building better understanding and models of jump dynamics.

Notable recent progress in this direction has been made by exploiting the bipower variation measure of the diffusive volatility component in the presence of jumps, introduced by Barndorff-Nielsen and Shephard (2004, 2006). The bipower variation, defined as the sum of products of adjacent absolute returns, has made it possible to disentangle the daily realized jump and diffusive volatility components (Huang and Tauchen, 2005; Andersen, Bollerslev, and Diebold, 2005; Andersen, Bollerslev, Frederiksen, and Nielsen, 2006) as well as to detect individual intraday jumps (Andersen, Bollerslev, and Dobrev, 2006; Andersen, Bollerslev, Frederiksen, and Nielsen, 2006). Very recently Christensen and Podolskij (2006b) have developed a potential improvement of the original bipower variation measure by summing products of adjacent ranges instead of absolute returns.

To the best of my knowledge, the closely related bipower variation measures of Barndorff-Nielsen and Shephard (2004, 2006) and Christensen and Podolskij (2006b) are the only existing alternatives to **GR** for estimating the diffusive volatility component in a jump-diffusion setting (with finite number of jumps on each day). In what follows, I carry out a realistic finite sample experiment and find that the efficiency of **GR** compares favorably to both existing alternatives.

4.1 Setup and Estimators

Based on the jump-diffusion models estimated on stock data in Andersen, Benzoni, and Lund (2002) and Eraker, Johannes, and Polson (2003), here I let the log-price process evolve according to a standard one-factor affine stochastic volatility model with Poisson jumps:

$$\begin{aligned} dp(t) &= \sigma(t)dW_1(t) + \kappa(t)dq(t), \\ d\sigma^2(t) &= \eta\left(\theta - \sigma^2(t)\right)dt + \nu\sigma(t)dW_2(t). \end{aligned} \tag{5}$$

The chosen stochastic volatility parameters are roughly consistent with the estimates in the above studies and have the following interpretation and values: daily unconditional mean $\theta = 1$, strength of mean reversion $\eta = 0.01$, volatility of volatility $\nu = 0.1$. The Poisson jump parameters are chosen to produce 20% contribution of jumps to total daily price variance on average, which is in line with the empirical findings in Andersen, Bollerslev, and Diebold (2005). In particular, the jump arrival rate is set to $\lambda = 1$, so that one jump per day arrives on average, while the jump size $\kappa(t)$ is distributed $N(0, 0.25)$.

Finally, a realistic "leverage effect" is incorporated by setting the correlation between return and volatility innovations to $\rho\left(dW_1(t), dW_2(t)\right) = -0.5$, which is also consistent with the empirically estimated levels on stock data. The resulting mean values of the daily total volatility and its breakdown in diffusive and jump components (in accordance with equation 3 above) are as follows: mean daily total volatility $\overline{QV} = \overline{IV} + \overline{JV} = 1.25$, mean daily diffusive volatility $\overline{IV} = 1$, and mean daily jump volatility $\overline{JV} = 0.25$.

In summary, the chosen simulation design is particularly realistic for stocks. Arguably, the results reported below are representative for a wide range of other reasonable specifications (available upon request).

The main goal here is to assess the finite sample efficiency of **GR**-based estimates of the daily diffusive volatility relative to the estimates based on bipower variation measures for realistic sample sizes n and number of price moves (returns) k taken by each estimator. In particular, I focus on the following four estimators of the daily IV:

1. The suitably scaled standard bipower variation measure $\mathbf{BV}_k$ based on k equidistant intraday returns as defined in Barndorff-Nielsen and Shephard (2004, 2006);

2. The suitably scaled subsample averaged bipower variation measure $\mathbf{SBV}_k$ obtained by averaging $\mathbf{BV}_k$ across all available non-overlapping subsamples of k intraday returns;[9]

3. The suitably scaled bipower range variation measure $\mathbf{RBV}_k$ based on k equidistant intraday ranges as defined in Christensen and Podolskij (2006b);

4. The suitably scaled generalized range measure $\mathbf{GR}_k$ based on the k largest non-overlapping intraday price moves as defined in this paper.

The proper finite-sample scaling factor for $\mathbf{BV}_k$ and $\mathbf{SBV}_k$ is $(\boldsymbol{\pi}/\mathbf{2}) \times k/(k-1)$, see Barndorff-Nielsen and Shephard (2004). The right finite sample scaling of $\mathbf{RBV}_k$ is achieved by Monte Carlo simulation based on the number of observations m falling within each intraday interval, as detailed in Christensen and Podolskij (2006b). Finally, the necessary finite sample scaling of $\mathbf{GR}_k$ is obtained by Monte Carlo simulation based on the daily sample size n in accordance with definition 6 above.

It is evident that $\mathbf{SBV}_k$, $\mathbf{RBV}_k$, and $\mathbf{GR}_k$ make use of all available data (although in a substantially different way), while $\mathbf{BV}_k$ is based only on k out of n observations.

[9] To the best of my knowledge, subsample averaging has not been formally considered yet in the context of bipower variations. However, following the insights of Zhang, Aït-Sahalia and Mykland (2005) and Aït-Sahalia, Mykland, and Zhang (2005b) in the context of realized volatility, it is natural to expect that subsample averaging is beneficial also for the bipower variation measure.

Therefore, it is meaningful to study the finite sample MSE of each estimator relative to the corresponding asymptotic variance of $\mathbf{BV}_k$ in the absence of jumps, whose convergence rate is $k^{1/2}$ (see Barndorff-Nielsen and Shephard, 2004). Given that they use more data, $\mathbf{SBV}_k$, $\mathbf{RBV}_k$, and $\mathbf{GR}_k$ should achieve somewhat higher efficiency than $\mathbf{BV}_k$ and, hence, smaller ratio for the relative MSE [10]

In order to make a finite sample comparison between these estimators for a wide realistic range of n and k, without much loss of generality it is convenient to restrict attention to values such that n is a multiple of k. In this case $m = n/k$ is exactly the number of subsamples across which $\mathbf{BV}_k$ is averaged to produce $\mathbf{SBV}_k$. In particular, I simulate data for 10,000 days of model (5) based on an Euler scheme and produce a representative panel of samples of two different types. The first sample type has a fixed daily sample size n = 1500, 3000, 7500, or 15000 with a corresponding set of twenty different values of k in the range from 2 to 500 that are divisors of each n. The second sample type keeps constant the ratio between n and k for the same values of k in the range from 2 to 500. In particular, $n_k = 5k$, $10k$, $25k$, or $100k$ is proportional to k, so that the number of available observations n_k is always $m = 5, 10, 25$ or 100 times larger than k.

Finally, the obtained 10,000 samples for each n or n_k are used to produce 10,000 corresponding daily diffusive volatility estimates based on each estimator and value of k. The efficiency of the estimators is then compared by producing two panels of volatility signature plots and RMSE signature plots for the two different sample types (see figures 3 and 4). The volatility signature plots show as a function of k the mean estimated diffusive volatility by each estimator across all days. The RMSE signature plots show as a function of k the RMSE of each estimator relative to the corresponding asymptotic variance of $\mathbf{BV}_k$ in the absence of jumps, having convergence rate $k^{1/2}$.

4.2 Discussion of Results

Figure 3a presents finite sample result for fixed sample sizes n = 1500, 3000, 7500, or 15000. The volatility signature plots show the mean of the obtained estimates across all days as a function of the number of returns k exploited by each estimator. The observed declining pattern of the upward bias of the estimators as k gets large is in accordance with their asymptotic unbiasedness and can be interpreted as a "jump signature plot". In particular, for very small k the estimators capture almost fully the total daily price variance (including the jump contribution) indicated by the upper bound on each graph,

[10] In particular, in the absence of jumps the ratio for $\mathbf{BV}_k$ should be about 1, while it should be smaller than 1 for the other three estimators. Importantly, in the presence of jumps all ratios should deteriorate (i.e. get larger) to a different extent depending on the jump-induced bias incurred by each estimator.

while for large k the mean estimates get closer to the diffusive part denoted by the lower bound at level 1. Importantly, **GR** appears to be the least biased estimator virtually for all considered n and $k \geq 10$.

Accordingly, the RMSE signature plots indicate that **GR** achieves the smallest RMSE offering modest efficiency gains compared to **RBV**, and more substantial gains compared to **SBV** and **BV**. The RMSE reduction **GR** vs. **BV** is typically in the range 50-60%, the RMSE reduction **GR** vs. **SBV** is typically in the range 40-50%, and the RMSE reduction **GR** vs. **RBV** is typically in the range 10-20%. Apparently, subsample averaging in this context does not lead to as much improvement as taking larger price moves. The RMSE plots suggest also that the larger the ratio between n and k, the more substantial is the benefit from taking the k largest moves.

To better check that this is indeed the case, figure 3b plots the results for sample sizes $n_k = 5k$, $10k$, $25k$ or $100k$ proportional to k. Indeed, while always observing the same relative standing of the estimators, the RMSE plots for **GR** are at different levels for each different ratio $m = n_k/k$, with relative RMSE just above 0.5 for ratio equal to 100 vs. relative RMSE about 0.7 for ratio equal to 5. In any case, the overall message from both figure 3a and 3b is the same: in the presence of jumps **GR** is a better estimator of the diffusive volatility component than its bipower variation alternatives.

Importantly, there is a simple explanation for the obtained conclusion. As detailed in the proof of proposition 7 above, what annihilates the distortion by jumps in the case of **GR** is the scaling factor e'_k shrinking to zero as k gets large regardless of whether jumps take place next to each other or apart from each other. By contrast, the bipower variation measures have a constant scaling factor and rely on the returns next to jumps to become small in the limit in order to annihilate the jump distortion. Of course, this does not always happen even in relatively large samples since consecutive jumps (or larger diffusive returns) can occur in arbitrary proximity to each other. This is one of the main reasons why the bipower variation measures are not as good in annihilating jumps as **GR**.

It is important to make sure, though, that the "jump signature plots" observed on figure 3a and 3b are not driven by some other source of bias than jumps, such as the simulated leverage effect. Therefore, figures 4a and 4b repeat the same plots for a scenario without jumps, i.e. when the total price variance is entirely due to diffusive volatility. As can be seen from these plots in the absence of jumps, all four volatility estimators have a mean very close to the true mean diffusive volatility equal to 1. Moreover, their relative standing in terms of RMSE is preserved also without jumps, although the efficiency gains are slightly lower in this case. This provides indirect evidence that only part of the efficiency gains of **GR** in the presence of jumps (figures 3a and 3b) are driven by its smaller

finite sample bias. There does not seem to be any other source of significant bias in this setting that could explain the documented superiority of **GR**.

Another important observation that can be made from the RMSE plots in figures 4a and 4b is that in the absence of jumps all four estimators seem to converge at the same asymptotic rate $k^{1/2}$. This is because they are unbiased and their RMSE stays flat relative to the asymptotic variance of **BV** in the absence of jumps, which is converging at the rate $k^{1/2}$. Then comparing back figures 4a and 4b to figures 3a and 3b one can see that jumps have a non-trivial impact on the efficiency of all four estimators as their RMSE increases by as much as 50%.

In summary, when it comes to measuring the pure diffusive volatility, **GR** looks promising in comparison to the other available alternatives. At the same time, improving the measurement of the diffusive volatility component is of particular help to recently developed jump detection procedures as in Andersen, Bollerslev, and Dobrev (2006) and volatility forecasting frameworks as in Andersen, Bollerslev, and Diebold (2005). The forecasting application in the following section sheds more light on this point.

5 Volatility Forecasting Based on GR

Andersen, Bollerslev, and Diebold (2005) have shown that almost all of the predictability in daily, weekly, and monthly return volatilities comes from the non-jump component of total price variance. Using **BV** as a robust to jumps volatility measure, the paper provides a practical framework for non-parametrically disentangling the jump and continuous components in asset return volatility. The method is applied to the DM/$ exchange rate, the S&P 500 market index, and the 30-year U.S. Treasury bond yield. In all these cases the results show that jumps are both highly prevalent and distinctly less persistent than the continuous sample path variation process.

Applying **GR** to the forecasting framework of Andersen, Bollerslev, and Diebold (2005) is a natural way to explore whether the modest efficiency gains from **GR**'s robustness to jumps lead to better volatility forecasting. As illustrated in the previous section on simulated data, **GR** has an advantage in measuring the pure diffusive component and compares well to **BV**.

In Andersen, Bollerslev, and Diebold (2005) days with price jumps are identified when the difference between **RV** and **BV** based on five minute returns is statistically significant according to the test of Barndorff-Nielsen and Shephard (2004, 2005). If a day has a jump than the jump component of volatility is represented by the difference between **RV** and **BV**, otherwise the jump component is zero. By analogy, to detect jump days based on

GR, I resort to the jump test developed in Andersen, Bollerslev, and Dobrev (2006). The jump component of volatility on days with detected jumps is measured by the difference between **RV** and **GR** or zero otherwise.

Volatility is forecasted based on the HAR-RV-CJ model of Andersen, Bollerslev, and Diebold (2005) for daily ($h = 1$), weekly ($h = 5$), and monthly ($h = 22$) volatility measurements:

$$\begin{aligned}\sqrt{RV_{t,t+h}} &= \beta_0 + \beta_{CD}\sqrt{C_t} + \beta_{CW}\sqrt{C_{t-5,t}} + \beta_{CM}\sqrt{C_{t-22,t}} \\ &\quad +\beta_{JD}\sqrt{J_t} + \beta_{JW}\sqrt{J_{t-5,t}} + \beta_{JM}\sqrt{J_{t-22,t}} + \varepsilon_{t,t+h},\end{aligned}$$

where RV is the average daily measure of total price variance in each period, while C and J are the corresponding average continuous and jump parts for a day, week, and a month.

I apply the HAR-RV-CJ forecasting framework to NYSE and NASDAQ traded stocks between July 2001 and December 2005 available through TAQ. I compare the forecast results with jump detection based on **GR** versus **BV**. For the sake of comparison, I focus on stocks traded on each day of the analyzed period.

When calculating GR on real stock data, it is important to limit the potential distortion caused by bid-ask bounce. This can be done by calculating GR directly on bid and ask quotes and exploiting its interpretation as the maximum ex-post trading gain or loss from k trades. In particular, each buy is based on an ask quote, while each sale is based on a bid quote. This leads to buying at local minima of the ask quotes and selling at local maxima of the bid quotes, as illustrated in the top panel of figure 5. The bid-ask spread automatically limits the maximum number of profitable price moves picked by GR, ensuring that it does not include moves below the spread level. Once the timings of all such price extrema are obtained, GR must be recalculated on the resultant mid quotes, as illustrated in the bottom panel of figure 5, assuming that on average mid-quotes provide unbiased estimates of the latent price. This two-step procedure limits the impact of microstructure noise on GR when calculated on real stock data.

Certain stocks are more prone to jumps and irregularities. Given NYSE and NASDAQ institutional differences, the NASDAQ traded stocks are more likely to exhibit jump volatility. In addition, the difference in the capitalization level is also an indicator of the potential level of jumps. I select both large and medium cap stocks. In this way, the results should be representative for different levels of jumps for liquid stocks. This investigation is limited to ten stocks on each exchange used for illustration purposes. To my knowledge, it is the first attempt to apply the above volatility forecasting framework to individual stocks.

Tables 1, 2 and 3 report the regression results for daily, weekly, and monthly volatility

forecasts respectively. The tables report the regression coefficients for each stock and the R-square from each regression. The results based on **GR** are close to those based on **BV**. In fact, **GR** consistently outperforms **BV** although by a small (but statistically significant) margin.[11] The regression coefficients confirm that the diffusive component is the main driver of the forecasting equation for individual stocks in close parallel with the findings in Andersen, Bollerslev, and Diebold (2005). However, the high R-square for stock indices that they report is not consistently observed for all stocks analyzed here, which is in line with generally lower predictability of volatility for individual stocks.

Nonetheless, this exercise provides indirect evidence that **GR** performs reasonably well as an estimator of the diffusive part of total price variance not only on simulations but also on real market data. The next section shows that even without imposing a particular model structure, **GR** proves useful for applications to relative value arbitrage.

6 Volatility-Based Rules for Pairs-Trading Using GR

Pairs trading is a relative value arbitrage rule which aims to identify a pair of stocks or other assets whose cumulative return processes are very likely to converge after a divergence. Upon spotting an instance of temporary relative mispricing traders try to make profits by going long the underpriced stock and shorting the overpriced stock in a zero investment portfolio. For understandable reasons the rules followed by traders to select and arbitrage such pairs of stocks are not disclosed.

Gatev, Goetzmann, and Rouwenhorst (2006) show that profitable pairs can be successfully identified by minimizing the distance between the cumulative return paths, i.e. normalized prices, across all possible stock pairs during a suitable formation period. Such pairs are then traded in the following months after the formation period by constructing a zero investment portfolio each time the prices have diverged by more than two historical standard deviations. The opened positions are unwinded whenever the prices cross again or at the end of the trading interval, whichever comes first. The resulting pairs-trading strategy generates significant excess returns that cannot be explained by conventional risk factors, including reversal. The main driver of the profits appears to be temporary mispricing of close substitutes.

Exploiting the fact that the profits are driven by the resulting convergence/divergence pattern, here I suggest two alternative volatility-based rules for choosing profitable stock pairs. The idea is to pick stock pairs whose price paths may not be as close in mean

[11] Both a sign test for the median and a t-test for the mean reject the hypothesis that the observed difference in R-square is zero.

square deviation terms but still converge and diverge. In essence, the proposed rules aim to recognize strong convergence/divergence patterns of higher amplitudes, i.e. pairs whose price difference has a zig-zag pattern with large ups and downs. For two stocks having price difference within the same range, the one with a wider zig-zag pattern will have higher volatility. Hence, minimizing the ratio of the range to realized volatility is one way to pick such stocks. Minimizing the ratio of the range to the generalized range is another even more direct way to find pairs of this kind given that the generalized range of order k is defined as the sum of the magnitudes of the largest k zig-zag moves on each path.

To put all this more formally, consider the price evolution of a pair of stocks S_t^A and S_t^B for a number of consecutive trading days $t = t_0, t_0 + 1, ..., t_0 + T$. Assume that on day t_0 a zero-investment portfolio is formed by buying \$1 worth of stock A and selling \$1 worth of stock B. On day $t_0 + T$ the portfolio is liquidated by unwinding both positions. For the sake of simplicity of the notation in this illustration, assume no stock splits and no dividends. Then the cumulative value of the portfolio on each day until liquidation will evolve according to:

$$P_t^{A,B} = \frac{1}{S_{t_0}^A} S_t^A - \frac{1}{S_{t_0}^B} S_t^B, \ t = t_0, t_0 + 1..., t_0 + T$$

In particular, $P_t^{A,B}$ is an excess return, $P_{t_0}^{A,B} = 0$, and $P_{t_0+T}^{A,B}$ is the generated profit upon liquidating the portfolio.

The choice of the stocks A and B is based on the historical evolution of the observed excess return $P_h^{A,B}$ during a prior formation period $h = 0, 1, ..., H$:

$$P_h^{A,B} = \frac{1}{S_0^A} S_h^A - \frac{1}{S_0^B} S_h^B, \ h = 0, 1..., H$$

In Gatev, Goetzmann, and Rouwenhorst (2006) the formation period is one year (about 250 business days) and the trading period is half year after the formation period (about 125 business days). Positions in pairs are opened each time their absolute excess return gets greater than two historical standard deviations, and pairs are chosen according to the following rule:

Rule 1 (SD): *For each stock A find another stock B, which minimizes the sum of squares* $SD^{A,B} = \sum_{h=0}^{H} \left[P_h^{A,B} \right]^2$ *of the cumulative excess returns* $P_h^{A,B}$, $h = 1, ..., H$ *observed during the formation period. Then take a set of pairs having the smallest* $SD^{A,B}$ *among all the obtained.*

The volatility-based rules that I suggest for finding stock pairs following a zig-zag pattern of wider amplitude are as follows:

Rule 2 (R/RV): *For each stock A find another stock B, which minimizes the ratio of the range to the realized volatility* $R^{A,B}/\sqrt{RV^{A,B}}$ *for the cumulative excess returns* $P_h^{A,B}$, $h = 1, ..., H$ *observed during the formation period, where* $R^{A,B} = \max_{h=0,...,H}\left[P_h^{A,B}\right] - \min_{h=0,...,H}\left[P_h^{A,B}\right]$ *and* $RV^{A,B} = \sum_{h=1}^{H}\left(P_h^{A,B} - P_{h-1}^{A,B}\right)^2$. *Then take a set of pairs having the smallest ratio* $R^{A,B}/\sqrt{RV^{A,B}}$ *among all the obtained.*

Rule 3 (R/GR): *For each stock A find another stock B, which minimizes the ratio of the range to the generalized range* $R^{A,B}/GR_k^{A,B}$ *for the cumulative excess returns* $P_h^{A,B}$, $h = 1, ..., H$ *observed during the formation period, where* $R^{A,B} = \max_{h=0,...,H}\left[P_h^{A,B}\right] - \min_{h=0,...,H}\left[P_h^{A,B}\right]$ *and* $GR_k = \max_{0\le h_1\le h_2\le ...\le h_{2k}\le H}\sum_{i=1}^{k}\left|P_{h_{2i}}^{A,B} - P_{h_{2i-1}}^{A,B}\right|$, *and k is the parameter that sets the number of large zig-zag moves*[12]. *Then take a set of pairs having the smallest ratio* $R^{A,B}/GR_k^{A,B}$ *among all the obtained.*

These rules pick largely disjoint sets of stock pairs characterized by different amplitudes of the zig-zag pattern followed by their normalized price difference. Figure 6 illustrates this point by plotting the normalized price paths of each stock and corresponding cumulative excess returns of the long-short portfolio for the 20th best pair obtained by the three rules. The formation period is 2003 and the trading period is the first half of 2004. The plotted stock pairs are representative for each rule regardless of the chosen formation period.

From figure 6 it is evident that the SD rule picks stocks whose normalized price paths are closest to each other and the corresponding excess returns for a \$1 long / \$1 short pair portfolio are tiny. The R/RV rule, and even more so the R/GR rule, pick stocks whose normalized prices diverge by more before converging, or equivalently, the resulting excess returns follow a zig-zag pattern of wider amplitudes. However, the trading threshold levels (say, two historical standard deviations) are also larger. Hence, from these graphs it is not possible to directly compare the potential gains from relative arbitrage.

A natural way to resolve this problem is to adjust the size of the initial long and short positions in each stock from \$1 to $\$g^{A,B}$, where the scaling factor $g^{A,B}$ for each pair is chosen to equate the corresponding trading threshold (say, two historical standard deviations as in Gatev, Goetzmann, and Rouwenhorst, 2006), to a common new level (say \$1):

[12]The larger the k, the closer rule 3 becomes to rule 2 in terms of chosen pairs. For k below 10 the rules have been found to produce largely disjoint sets of best 50 stock pairs.

$$g^{A,B} = \frac{1}{2\sigma\left[P_h^{A,B}\right]}$$

Clearly, with this choice

$$2\sigma\left[g^{A,B}\frac{1}{S_0^A}S_h^A - g^{A,B}\frac{1}{S_0^B}S_h^B\right] = 2\sigma\left[g^{A,B}P_h^{A,B}\right] = 1\ ,$$

which ensures that a relative arbitrage position in each pair scaled by $g^{A,B}$ should be opened exactly when the historical excess return path of the same scaled position hits the \$1 or -\$1 levels.

Figure 7 plots the scaled and unscaled excess return paths of the 50 best pairs obtained by each rule for the same formation and trading periods as above (2003 and the first half of 2004). Plotting multiple pairs better reveals the overall characteristics of the stock pairs picked by each rule. The left column of figure 7 shows the unscaled excess returns as in the right column of figure 6, while the right column scales up the positions in each stock pair to a common trading threshold level of ±\$1. The mean scaling factor for the selected pairs is reported in the upper left corner of each scaled plot. One thing to notice is that the scaling is largest for the SD rule and smallest for the R/GR rule. Since excess returns with smaller zig-zag amplitudes need to be scaled by more, the scaling factor g directly indicates the relative amplitude of the zig-zag. The obtained ordering of the rules with respect to scaling is robust to the formation period and the number of stocks chosen (the fewer the pairs, the larger the difference). The results are in the same spirit for other possible scalings that equalize the range or the realized volatility of the historical excess returns instead of the trading thresholds.

It is worth noting that by means of scaling each rule can be gauged to yield potentially the same gains from relative value arbitrage. However, if there is a common limit to scaling, then the R/GR rule will be most preferable since it selects stock pairs of widest zig-zag pattern, which require the smallest scaling of the relative arbitrage trades[13].

Since the three sets of stock pairs picked by the different rules are largely disjoint, it is not immediately clear whether the resulting excess returns from pairs trading based on the R/RV and R/GR rules will be driven by the same factors as those based on the SD

[13]The larger is the dollar value of the executed trades, the higher is the price pressure in the opposite direction that undermines the profitability of the strategy. Another capacity constraint could arise from limits to short selling.

rule documented by Gatev, Goetzmann, and Rouwenhorst (2006). Further investigation is necessary to reveal the driving risk factors and to shed light on the observed more persistent pattern from the formation into the trading period. The initial findings in concurrent work in progress suggest that the volatility-based strategies are profitable on average and require fewer trades at the expense of higher risk, especially without filtering out penny stocks.

In conclusion, as in Gatev, Goetzmann, and Rouwenhorst (2006), the proposed volatility-based rules are able to pick stock pairs suitable for relative value arbitrage. The selected stock pairs are characterized by larger amplitudes of the zig-zag pattern followed by the cumulative excess returns of the corresponding arbitrage portfolio. This pattern is robust to the pair formation period and carries well into the trading period.

7 Concluding Remarks

This paper has introduced a novel high-frequency volatility estimator based on large price moves, which constitutes a generalization of the range. Just like the standard range (high minus low) maximizes a single price difference on the observed price path, the generalized range maximizes the sum of multiple price differences (between consecutive peaks and troughs). It has an intuitive interpretation as the maximum ex-post gain or loss from multiple non-overlapping intraday trades. Unlike volatility estimators that take multiple intraday returns or ranges on a fixed time grid, GR involves maximization on a time grid. In this way, GR provides a new link between recent advances in volatility estimation on high-frequency data with the long-established efficiency of the daily range as a volatility estimator.

The main contribution of this paper is the derivation of the almost sure asymptotics of GR in a jump-diffusion setting. In particular, when properly scaled, GR defines novel strongly consistent and asymptotically unbiased estimators of the diffusive component of total volatility given that jumps of finite activity are annihilated by the required scaling factor. Building the underlying theory marks a departure from the well-established theoretical framework of realized volatility since GR is driven by extreme points on the observed price path without fixing a time grid. In order to derive the almost sure asymptotics of GR, I utilize distributional properties of the Brownian h-extrema, which is another contribution to the volatility literature as a useful new tool for analysis.

The documented performance on simulated data from a realistic jump-diffusion shows that, in accordance with the intuition, GR compares favorably to the other known type of diffusive volatility estimators that are robust to jumps, i.e. the bipower variation measures.

This is partly because in finite samples GR annihilates jumps in a more robust way than the bipower measures. A volatility forecasting application on stock data, which exploits the separation of the jump and the diffusive components of volatility, confirms the efficiency gains from using GR documented on simulated data. In an application to relative value arbitrage strategies, the capability of GR to identify large zig-zag price moves appears to be valuable even without making particular assumptions on the dynamics of the price process or the sampling scheme.

By introducing the novel generalized range measure of volatility and its unique properties, this paper sets the stage for research in various application areas and further theoretical advancements. On the theory front, the Brownian h-extrema used to derive the core result in this paper, clearly call for further investigation as they open an interesting possibility to estimate volatility based on the passage times between price peaks and troughs in a new duration-based approach to volatility estimation. Another more direct, though difficult, theory extension would be deriving the limiting distribution of GR in the absence of jumps.

On the application side, first, the framework of Brandt and Diebold (2006) for covariance measurement based on the standard range can be readily extended to the generalized range. In essence, this would allow measuring asset comovements and market betas from large price moves. Second, as suggested by the pairs trading application, GR appears to be directly applicable to arbitrage strategies, market timing, and volatility trading. Third, the ability of GR to identify jumps and large price moves can help studying their determinants from an information perspective. Likewise, studying the limiting behavior of GR in the presence of microstructure noise could shed more light on the dynamics of bid and ask quotes within the modeling framework of Hasbrouck (1999). Last, by definition, GR can be used to define a natural ex-post measure of the market timing ability of traders or their risk exposure, extending the maximum drawdown measure of risk used by practitioners. Overall, GR opens a number of intriguing venues for future research in the areas of risk management, portfolio choice, and market microstructure.

A Appendix: Proofs

Proof of property 1. The equality $\mathbf{GR}_k(\mathbf{b}_{[0,\boldsymbol{\tau}(T)]}) = \sqrt{IV_{[0,T]}} \times \mathbf{GR}_k(\mathbf{B}_{[0,1]})$ is established by the relationship

$$\begin{aligned}\sum_{i=1}^{k} |\mathbf{b}(\boldsymbol{\tau}(t_{2i})) - \mathbf{b}(\boldsymbol{\tau}(t_{2i-1}))| &= \sqrt{\boldsymbol{\tau}(T)} \times \sum_{i=1}^{k} \left| \frac{\mathbf{b}(\boldsymbol{\tau}(t_{2i}))}{\sqrt{\boldsymbol{\tau}(T)}} - \frac{\mathbf{b}(\boldsymbol{\tau}(t_{2i-1}))}{\sqrt{\boldsymbol{\tau}(T)}} \right| \\ &= \sqrt{\boldsymbol{\tau}(T)} \times \sum_{i=1}^{k} \left| \mathbf{B}\left(\frac{\boldsymbol{\tau}(t_{2i})}{\boldsymbol{\tau}(T)} \right) - \mathbf{B}\left(\frac{\boldsymbol{\tau}(t_{2i-1})}{\boldsymbol{\tau}(T)} \right) \right|\end{aligned}$$

where $\boldsymbol{\tau}(T) = \sqrt{IV_{[0,T]}}$. The only extra thing to notice is that $\mathbf{B}(z) = \frac{1}{\sqrt{\boldsymbol{\tau}(T)}} \mathbf{b}(z \times \boldsymbol{\tau}(T))$ is the rescaled version of the Brownian motion $\mathbf{b}$ mapping the interval $[0, \boldsymbol{\tau}(T)]$ for $\boldsymbol{\tau}$ to the unit interval $[0,1]$ for z, so by the scaling property of Brownian motion it follows that $\mathbf{B}_{[0,1]}$ is some standard Brownian path on $[0,1]$. ■

Lemma 9 *Let Ω be the Brownian path space, i.e. the space of continuous functions $\omega : [0,\infty] \longrightarrow \mathbb{R}$ equipped with the Wiener measure. For each path ω let $g_k(\omega)$ be a threshold level and $q_k(\omega)$ be a corresponding integer between 0 and k for which $\mathbf{GR}_k^{\ominus}$ calculated on the path ω is given by the sum of the largest k among $k+q_k$ absolute differences between $k+q_k+1$ Brownian g_k-extrema in the interval $[0,1]$. Then the random sequences $\{q_k(\omega)\}_{k=1}^{\infty}$ and $\{g_k(\omega)\}_{k=1}^{\infty}$ have the following limiting property:*

$$\limsup_{k \longrightarrow \infty} (k+q_k)^{1/2} g_k \leq 1 \ a.s.$$

Proof. Consider the durations δ_t^k, $t = 1,\ 2\ ,\ ...,\ k+q_k$ between the $k+q_k+1$ Brownian g_k-extrema on each path ω. Since all of the extrema fall in the interval $[0,1]$, the sum of these durations cannot exceed 1:

$$\sum_{t=1}^{k+q_k} \delta_t^k \leq 1 \text{ for each } k$$

At the same time, the characterization of Brownian h-extrema by Neveu and Pitman (1989) implies that for each k the durations δ_t^k, $t = 1,\ 2\ ,\ ...,\ k+q_k$ are some i.i.d. draw of random variables with mean g_k^2. Hence, the sum of these durations has mean $(k+q_k)\, g_k^2$ and

$$\sum_{t=1}^{k+q_k} \delta_t^k - (k+q_k)\, g_k^2 \xrightarrow{a.s.} 0$$

Therefore, for any $\varepsilon > 0$

$$P\left\{\omega : (k+q_k)\, g_k^2 - \sum_{t=1}^{k+q_k} \delta_t^k > \varepsilon, \text{ i.o.}\right\} = 0$$

Hence,

$$P\left\{\omega : \sum_{t=1}^{k+q_k} \delta_t^k \leq 1 \text{ and } (k+q_k)\, g_k^2 > 1+\varepsilon, \text{ i.o.}\right\} = 0$$

which in turn implies

$$\begin{array}{rcl} \limsup_{k\longrightarrow\infty} (k+q_k)\, g_k^2 & \leq & 1 \ a.s. \\ & \Updownarrow & \\ \limsup_{k\longrightarrow\infty} (k+q_k)^{1/2}\, g_k & \leq & 1 \ a.s. \end{array}$$

Q.E.D. ■

Lemma 10 *Let Ω be the Brownian path space, i.e. the space of continuous functions $\omega : [0,\infty] \longrightarrow \mathbb{R}$ equipped with the Wiener measure. Let $\{q_k(\omega)\}_{k=1}^{\infty}$ be a random sequence of non-negative integers such that q_k/k is uniformly bounded a.s. and $\{g_k(\omega)\}_{k=1}^{\infty}$ be a random sequence of real positive numbers on Ω such that*

$$\limsup_{k\longrightarrow\infty} (k+q_k)^{1/2}\, g_k \leq 1 \ a.s.$$

Given any such g_k and q_k, take $k+q_k+1$ alternating Brownian g_k-extrema on each Brownian path and let $E_{(1)}^k(\omega)$, $E_{(2)}^k(\omega)$, ..., $E_{(k+q_k)}^k(\omega)$ be the corresponding ascendingly ordered statistics of absolute differences between adjacent g_k-extrema. Finally, define

$$G_k(\omega) \doteq \sum_{t=1}^{k} E_{(t+q_k)}^k(\omega)$$

as the sum of the largest k order statistics $E_{(1+q_k)}^k(\omega), E_{(2+q_k)}^k(\omega)$, ..., $E_{(k+q_k)}^k(\omega)$. Then

$$\limsup_{k\longrightarrow\infty} \frac{1}{2} k^{-\frac{1}{2}} G_k \leq 1 \ a.s.$$

Proof. From the characterization of Brownian h-extrema by Neveu and Pitman (1989), it follows that $E_{(1)}^k - g_k$, $E_{(2)}^k - g_k$, ..., $E_{(k+q_k)}^k - g_k$ are the exponential order statistics of

some i.i.d. draw from an exponential distribution with mean g_k. On the other hand, it is a known fact that the exponential order statistics have independent and exponentially distributed differences (see Feller, 1971), implying in particular:

$$E^k_{(i)} - g_k = \sum_{j=1}^{i} \frac{Z^k_j}{k + q_k + 1 - j}$$

where the variables Z^k_j are an i.i.d. draw from an exponential distribution with mean g_k. In this way,

$$\begin{aligned}
\frac{1}{2}k^{-\frac{1}{2}}G_k &= \frac{1}{2}k^{-\frac{1}{2}}\sum_{t=1}^{k} E^k_{(t+q_k)} \\
&= \frac{1}{2}k^{-\frac{1}{2}}\sum_{t=1}^{k}\sum_{j=1}^{t+q_k}\left(\frac{Z^k_j}{k+q_k+1-j} + g_k\right) \\
&= \frac{1}{2}k^{-\frac{1}{2}}\left(kg_k + \sum_{t=1}^{k}\sum_{j=1}^{q_k}\frac{Z^k_j}{k+q_k+1-j} + \sum_{t=1}^{k}\sum_{j=1+q_k}^{t+q_k}\frac{Z^k_j}{k+q_k+1-j}\right) \\
&= \frac{1}{2}k^{-\frac{1}{2}}\left(kg_k + k\sum_{j=1}^{q_k}\frac{Z^k_j}{k+q_k+1-j} + \sum_{j=1+q_k}^{k+q_k} Z^k_j\right) \\
&= \frac{1}{2}k^{-\frac{1}{2}}g_k\left(k + k\sum_{j=1}^{q_k}\frac{Z^k_j/g_k}{k+q_k+1-j} + \sum_{j=1+q_k}^{k+q_k} Z^k_j/g_k\right) =
\end{aligned}$$

Note that for each k the scaled variables Z^k_j/g_k are i.i.d. exponential with mean 1. Hence, it is useful to rearrange the last expression as:

$$\begin{aligned}
1/2k^{-1/2}G_k &= \frac{1}{2}k^{-\frac{1}{2}}g_k\left(k + k\sum_{j=1}^{q_k}\frac{1}{k+q_k+1-j} + \sum_{j=1+q_k}^{k+q_k} 1\right) \\
&\quad + \frac{1}{2}k^{-\frac{1}{2}}g_k\left(\sum_{j=1}^{q_k}\frac{k}{k+q_k+1-j}\left(Z^k_j/g_k - 1\right) + \sum_{j=1+q_k}^{k+q_k}\left(Z^k_j/g_k - 1\right)\right)
\end{aligned}$$

This representation is of the form

$$\frac{1}{2}k^{-\frac{1}{2}}G_k = D_k + R_k$$

where the first part is

$$D_k = \frac{1}{2} k^{-\frac{1}{2}} g_k \left(2k + k \sum_{j=1}^{q_k} \frac{1}{k + q_k + 1 - j} \right)$$

while the second part is

$$R_k = \frac{1}{2} k^{-\frac{1}{2}} g_k \left(\sum_{j=1}^{q_k} \frac{k}{k + q_k + 1 - j} \left(Z_j^k / g_k - 1 \right) + \sum_{j=1+q_k}^{k+q_k} \left(Z_j^k / g_k - 1 \right) \right)$$

Taking into account the structure of these expressions, it suffices to show that, first, $R_k \xrightarrow{a.s.} 0$ and, second, $\limsup_{k \longrightarrow \infty} D_k \leq 1$ *a.s.*.

To deal with R_k note that

$$\frac{1}{k + q_k} \left(\sum_{j=1}^{q_k} \frac{k}{k + q_k + 1 - j} \left(Z_j^k / g_k - 1 \right) + \sum_{j=1+q_k}^{k+q_k} \left(Z_j^k / g_k - 1 \right) \right) \xrightarrow{a.s.} 0$$

given that $Z_j^k / g_k - 1,\ j = 1, 2, ..., k + q_k$ are $k + q_k$ i.i.d. variables with mean 0 and variance 1 and the factors in front of each of them in the summation are strictly bounded away from 0 and do not exceed 1. The limit is preserved after multiplication by $1/2k^{-1/2} g_k \left(k + q_k \right)$ because

$$\frac{1}{2} k^{-\frac{1}{2}} g_k \left(k + q_k \right) = \frac{1}{2} \left[\left(k + q_k \right)^{1/2} g_k \right] \left[\left(1 + q_k / k \right)^{1/2} \right]$$

is uniformly bounded a.s., given that by assumption both q_k / k and $\left(k + q_k \right)^{1/2} g_k$ are uniformly bounded a.s. Thus, $R_k \xrightarrow{a.s.} 0$ as desired.

As far as D_k is concerned, it can be further manipulated as follows:

$$\begin{aligned} D_k &= \frac{1}{2} k^{1/2} g_k \left(2 + \sum_{j=1}^{q_k} \frac{1}{k + q_k + 1 - j} \right) \\ &= \left[\left(k + q_k \right)^{1/2} g_k \right] \left[\left(1 + q_k / k \right)^{-1/2} \frac{1}{2} \left(2 + \sum_{j=1}^{q_k} \frac{1}{k + j} \right) \right] \end{aligned}$$

Note that by assumption

$$\limsup_{k \longrightarrow \infty} \left[\left(k + q_k \right)^{1/2} g_k \right] \leq 1 \ a.s.$$

while it is straightforward to show by induction on $q_k = 0, 1, 2, ...$ that

$$\left[(1 + q_k/k)^{-1/2} \frac{1}{2} \left(2 + \sum_{j=1}^{q_k} \frac{1}{k+j} \right) \right] \leq 1 \text{ for all } q_k \geq 0$$

Therefore, $\limsup_{k \longrightarrow \infty} D_k \leq 1$ $a.s.$ Combining the obtained results for D_k and R_k yields

$$\limsup_{k \longrightarrow \infty} \frac{1}{2} k^{-\frac{1}{2}} G_k = \limsup_{k \longrightarrow \infty} (D_k + R_k) \leq 1 \ a.s.$$

Q.E.D. ■

Corollary 11 $\limsup_{k \longrightarrow \infty} \frac{1}{2} k^{-\frac{1}{2}} \mathbf{GR}_k^{\ominus} \leq 1$ $a.s.$

Lemma 12 $\liminf_{k \longrightarrow \infty} \frac{1}{2} k^{-\frac{1}{2}} \mathbf{GR}_k^{\ominus} \geq 1$ $a.s.$

Proof. On each Brownian path $\mathbf{B}$ (not anymore restricted to the unit interval) define the h-extremal generalized range $\mathbf{HGR}_k[h]$ of order k and threshold h as the sum of the k price moves between $k+1$ consecutive h-extrema $\mathbf{B}(s_{0,[h]})$, $\mathbf{B}(s_{1,[h]})$, $\mathbf{B}(s_{k,[h]})$ on this path, i.e.

$$\mathbf{HGR}_{k,[h]} \doteq \sum_{j=1}^{k} \left| \mathbf{B}(s_{j,[h]}) - \mathbf{B}(s_{j-1,[h]}) \right|$$

Note that the corresponding total duration (or interval length) $\Delta_{k,[h]} = s_{k,[h]} - s_{0,[h]}$ is randomly varying with each path. However, the first two moments of both $\mathbf{HGR}_{k,[h]}$ and its total duration $\Delta_{k,[h]}$ easily follow from Neveu's and Pitman's (1989) characterization of Brownian h-extrema and their times:

$$\begin{aligned} \mathbb{E}\left[\mathbf{HGR}_{k,[h]}\right] &= 2kh, \, \mathbb{E}\left[\Delta_{k,[h]}\right] = kh^2, \ \ k = 1, 2, ... \\ \mathbb{V}\left[\mathbf{HGR}_{k,[h]}\right] &= kh^2, \, \mathbb{V}\left[\Delta_{k,[h]}\right] = kh^4, \ \ k = 1, 2, ... \end{aligned}$$

In particular, if the threshold level h is set to $h_k(\varepsilon) = (1-\varepsilon) k^{-1/2}$ for any fixed $0 < \varepsilon < 1$, it follows that the expected value of $\mathbf{HGR}_{k,[h_k(\varepsilon)]}$ is $2k^{1/2}(1-\varepsilon)$ with variance $(1-\varepsilon)^2$, while the expected value of its duration $\Delta_{k,[h_k(\varepsilon)]}$ is $(1-\varepsilon)^2$ with variance $(1-\varepsilon)^4/k$. Moreover, by independence it follows:

$$\begin{aligned} &\Delta_{k,[h_k(\varepsilon)]} \xrightarrow{a.s.} (1-\varepsilon)^2 \\ &\frac{1}{2} k^{-\frac{1}{2}} \mathbf{HGR}_{k,[h_k(\varepsilon)]} \xrightarrow{a.s.} (1-\varepsilon) \end{aligned}$$

Consequently,

$$P\left\{\omega : \Delta_{k,[h_k(\varepsilon)]} > 1, \text{ i.o.}\right\} = 0$$

and

$$P\left\{\omega : \frac{1}{2}k^{-\frac{1}{2}}\mathbf{HGR}_{k,[h_k(\varepsilon)]} < 1 - 2\varepsilon, \text{ i.o.}\right\} = 0$$

On the other hand, as long as $\Delta_{k,[h_k(\varepsilon)]} \leq 1$, extremality of $\mathbf{GR}_k$ in $[0,1]$ guarantees that

$$\mathbf{HGR}_{k,[h_k(\varepsilon)]} \leq \mathbf{GR}_k^{\ominus}$$

Hence,

$$P\left\{\omega : \frac{1}{2}k^{-\frac{1}{2}}\mathbf{GR}_k^{\ominus} < 1 - 2\varepsilon, \text{ i.o.}\right\} = 0$$

Since this is true for an arbitrary $0 < \varepsilon < 1$, it follows

$$\liminf_{k\longrightarrow\infty} \frac{1}{2}k^{-\frac{1}{2}}\mathbf{GR}_k^{\ominus} \geq 1 \ a.s.$$

Q.E.D. ■

Proof of proposition 4. Let $\varepsilon > 0$ be arbitrary small. To put a suitable bound on the discrepancy between $\mathbf{GR}_k^n$ and $\mathbf{GR}_k^{\ominus}$, observe that Levy's modulus of continuity implies that there exist $\delta > 0$ such that a.s. $|\mathbf{B}_{t+s} - \mathbf{B}_t| < 2\left(s\left|\log s\right|\right)^{1/2}$ for any $t \in [0,1]$ whenever $s < \delta$. Now let $\overline{\Omega} \subseteq \Omega$ be the set of Brownian paths ω with $P\left(\overline{\Omega}\right) = 1$ on which this inequality holds and the a.s. limit $1/2k^{-1/2}\mathbf{GR}_k^{\ominus} \longrightarrow 1$ established above is also valid.

Take any path $\omega \in \overline{\Omega}$. It is given that $k\Delta_n\left|\log\Delta_n\right| \longrightarrow 0$, hence, one can find a number $l(\omega,\epsilon)$ such that

$$\left(k\Delta_n\left|\log\Delta_n\right|\right)^{1/2} < \varepsilon/4 \text{ for all } k > l(\omega,\epsilon).$$

In addition, it is possible to find a number $m(\omega,\epsilon)$ such that

$$\left|1/2k^{-1/2}\mathbf{GR}_k^{\ominus} - 1\right| < \varepsilon/2 \text{ for all } k > m(\omega,\epsilon).$$

Finally, since $k\left(\Delta_n\left|\log\Delta_n\right|\right)^{1/2} \longrightarrow 0$ implies $\Delta_n \longrightarrow 0$, there is a number $p(\omega,\delta)$ such that $\Delta_n < \delta$ for all $k > p(\omega,\delta)$.

Define $k(\omega,\epsilon) = \max\left\{l(\omega,\epsilon),\ m(\omega,\epsilon),\ p(\omega,\delta)\right\}$ and observe that the discrepancy be-

tween $\mathbf{GR}_k^n$ and $\mathbf{GR}_k^{\ominus}$ is bounded by

$$2k \sup_{0\leq t\leq 1} |\mathbf{B}_{t+\Delta_n} - \mathbf{B}_t|$$

since $\mathbf{GR}_k^n$ and $\mathbf{GR}_k^{\ominus}$ are calculated on at most $2k$ Brownian points (alternating local extrema) and the mesh size is Δ_n. By the above choice of $k(\omega,\epsilon)$ it follows that $\Delta_n < \delta$ whenever $k > k(\omega,\epsilon)$, which in turn implies

$$\left|\mathbf{GR}_k^n - \mathbf{GR}_k^{\ominus}\right| < 4k\left(\Delta_n \left|\log \Delta_n\right|\right)^{1/2} = k^{1/2} 4 \left(k\Delta_n \left|\log \Delta_n\right|\right)^{1/2} < k^{1/2}\varepsilon$$

Therefore, the following chain of inequalities holds for all $k > k(\omega,\epsilon)$:

$$\begin{aligned}
\left|1/2k^{-1/2}\mathbf{GR}_k^n - 1\right| &\leq \left|1/2k^{-1/2}\mathbf{GR}_k^n - 1/2k^{-1/2}\mathbf{GR}_k^{\ominus}\right| + \left|1/2k^{-1/2}\mathbf{GR}_k^{\ominus} - 1\right| \\
&= 1/2k^{-1/2}\left|\mathbf{GR}_k^n - \mathbf{GR}_k^{\ominus}\right| + \left|1/2k^{-1/2}\mathbf{GR}_k^{\ominus} - 1\right| \\
&\leq \epsilon/2 + \epsilon/2 = \epsilon
\end{aligned}$$

This concludes the proof as we have shown that $\left|1/2k^{-1/2}\mathbf{GR}_k^n - 1\right| < \epsilon$ whenever $k > k(\omega,\epsilon)$ for any Brownian path $\omega \in \overline{\Omega}$, where $P\left(\overline{\Omega}\right) = 1$. ■

Proof of proposition 7. It suffices to demonstrate strong consistency of the first of the three estimators $e_k'\mathbf{GR}_k(\mathbf{p}_{[0,T]})$ as the argument for the other two is along the same lines. In the absence of jumps and drift, the strong consistency result

$$e_k'\mathbf{GR}_k(\mathbf{p}_{[0,T]}) \xrightarrow{a.s.} \sqrt{IV_{[0,T]}}$$

follows immediately by combining variance separability of **GR** on a Brownian diffusion (property P1) with the almost sure asymptotics of **GR** on a standard Brownian motion (propositions 2 and 4).

Thus, it remains to show that a finite number of jumps does not change the asymptotic limit of the product $e_k'\mathbf{GR}_k(\mathbf{p}_{[0,T]})$. To this end, in the presence of jumps it is convenient to express $\mathbf{p}$ as the sum of its continuous diffusive part $\mathbf{p}^c$ and the complementing jump part $\mathbf{p}^j$:

$$\mathbf{p}_t = \mathbf{p}_t^c + \mathbf{p}_t^j \ , t \in [0,T]$$

Then if $0 \leq t_1^* \leq t_2^* \leq \cdots \leq t_{2k}^* \leq T$ denote the optimal times for $\mathbf{GR}_k(\mathbf{p}_{[0,T]})$, i.e. $\sum_{i=1}^{k}\left|\mathbf{p}(t_{2i}^*) - \mathbf{p}(t_{2i-1}^*)\right| \equiv \mathbf{GR}_k(\mathbf{p}_{[0,T]})$, the definition of **GR** and the triangle inequality

imply:

$$\begin{aligned}
\mathbf{GR}_k(\mathbf{p}_{[0,T]}) &= \sum_{i=1}^{k} \left|\mathbf{p}(t_{2i}^*) - \mathbf{p}(t_{2i-1}^*)\right| \\
&= \sum_{i=1}^{k} \left|\mathbf{p}^c(t_{2i}^*) - \mathbf{p}^c(t_{2i-1}^*) + \mathbf{p}^j(t_{2i}^*) - \mathbf{p}^j(t_{2i-1}^*)\right| \\
&\leq \sum_{i=1}^{k} \left|\mathbf{p}^c(t_{2i}^*) - \mathbf{p}^c(t_{2i-1}^*)\right| + \sum_{i=1}^{k} \left|\mathbf{p}^j(t_{2i}^*) - \mathbf{p}^j(t_{2i-1}^*)\right| \\
&\leq \mathbf{GR}_k(\mathbf{p}_{[0,T]}^c) + \mathbf{GR}_k(\mathbf{p}_{[0,T]}^j).
\end{aligned}$$

Likewise, if $0 \leq t_1^\circ \leq t_2^\circ \leq \cdots \leq t_{2k}^\circ \leq T$ denote the optimal times for $\mathbf{GR}_k(\mathbf{p}_{[0,T]}^c)$ calculated on the continuous part of $\mathbf{p}$, then

$$\begin{aligned}
\mathbf{GR}_k(\mathbf{p}_{[0,T]}) &\geq \sum_{i=1}^{k} \left|\mathbf{p}(t_{2i}^\circ) - \mathbf{p}(t_{2i-1}^\circ)\right| \\
&= \sum_{i=1}^{k} \left|\mathbf{p}^c(t_{2i}^\circ) - \mathbf{p}^c(t_{2i-1}^\circ) + \mathbf{p}^j(t_{2i}^\circ) - \mathbf{p}^j(t_{2i-1}^\circ)\right| \\
&\geq \sum_{i=1}^{k} \left|\mathbf{p}^c(t_{2i}^\circ) - \mathbf{p}^c(t_{2i-1}^\circ)\right| - \sum_{i=1}^{k} \left|\mathbf{p}^j(t_{2i}^\circ) - \mathbf{p}^j(t_{2i-1}^\circ)\right| \\
&= \mathbf{GR}_k(\mathbf{p}_{[0,T]}^c) - \sum_{i=1}^{k} \left|\mathbf{p}^j(t_{2i}^\circ) - \mathbf{p}^j(t_{2i-1}^\circ)\right| \\
&\geq \mathbf{GR}_k(\mathbf{p}_{[0,T]}^c) - \mathbf{GR}_k(\mathbf{p}_{[0,T]}^j).
\end{aligned}$$

Combining the above two chains of inequalities yields

$$\left|\mathbf{GR}_k(\mathbf{p}_{[0,T]}) - \mathbf{GR}_k(\mathbf{p}_{[0,T]}^c)\right| \leq \mathbf{GR}_k(\mathbf{p}_{[0,T]}^j).$$

Here by definition the jump component is the sum $\mathbf{p}_{[0,T]}^j \equiv \sum_{0<s\leq T} \kappa(s)$. Therefore, as long as k gets larger than the corresponding finite number of jumps in $[0,T]$ it must be the case that

$$\mathbf{GR}_k(\mathbf{p}_{[0,T]}^j) \equiv \sum_{0<s\leq T} |\kappa(s)| < \infty.$$

On the other hand, the scaling factor e_k' shrinks to zero at a rate $k^{-1/2}$ implying in

particular that

$$e_k' \mathbf{GR}_k(\mathbf{p}^j_{[0,T]}) = e_k' \sum_{0<s\leq T} |\kappa(s)| \longrightarrow 0.$$

Hence,

$$\left| e_k' \mathbf{GR}_k(\mathbf{p}_{[0,T]}) - e_k' \mathbf{GR}_k(\mathbf{p}^c_{[0,T]}) \right| \longrightarrow 0.$$

Finally, the already established a.s. limit in the absence of jumps implies that for the continuous part of $\mathbf{p}$

$$e_k' \mathbf{GR}_k(\mathbf{p}^c_{[0,T]}) \xrightarrow{a.s.} \sqrt{IV_{[0,T]}},$$

which combined with the preceding asymptotic equivalence yields the same asymptotic limit without excluding the jumps from $\mathbf{p}$:

$$e_k' \mathbf{GR}_k(\mathbf{p}_{[0,T]}) \xrightarrow{a.s.} \sqrt{IV_{[0,T]}}$$

More generally, the same derivation would hold also after adding any process of bounded variation. In particular, it follows that the obtained asymptotic limit would remain valid in the presence of a non-zero drift in the jump-diffusion sde (2).

Q.E.D. ■

Proof of proposition 8. Again, it suffices to show asymptotic unbiasedness of the first of the three estimators $e_k' \mathbf{GR}_k(\mathbf{p}_{[0,T]})$ as the argument for the other two is analogous.

It is convenient first to consider the case without jumps and drift, and then use the same argument as in the proof of proposition 7 above to show that the impact of a finite number of jumps or a bounded variation drift is asymptotically negligible.

In the absence of jumps and drift, by variance separability (property P1)

$$\mathbf{GR}_k(\mathbf{p}_{[0,T]}) = \sqrt{IV_{[0,T]}} \times \mathbf{GR}_k(\mathbf{B}_{[0,1]})$$

on every path. Hence, given that $e_k' = \left(\mathbb{E}\left[\mathbf{GR}_k(\mathbf{B}_{[0,1]})\right]\right)^{-1}$ it follows:

$$\begin{aligned} \mathbb{E}\left[e_k' \mathbf{GR}_k(\mathbf{p}_{[0,T]}) \mid IV_{[0,T]}\right] &= e_k' \sqrt{IV_{[0,T]}} \mathbb{E}\left[\mathbf{GR}_k(\mathbf{B}_{[0,1]}) \mid IV_{[0,T]}\right] \\ &= \sqrt{IV_{[0,T]}} \frac{\mathbb{E}\left[\mathbf{GR}_k(\mathbf{B}_{[0,1]}) \mid IV_{[0,T]}\right]}{\mathbb{E}\left[\mathbf{GR}_k(\mathbf{B}_{[0,1]})\right]} \end{aligned}$$

Now, if the time change is independent, i.e. there is no "leverage effect", then the value of

$IV_{[0,T]}$ does not contain any information about the corresponding set of Brownian paths $\mathbf{B}_{[0,1]}$ and, hence, $\mathbb{E}\left[\mathbf{GR}_k(\mathbf{B}_{[0,1]}) \mid IV_{[0,T]}\right] \equiv \mathbb{E}\left[\mathbf{GR}_k(\mathbf{B}_{[0,1]})\right]$. However, in case the value of $IV_{[0,T]}$ restricts the attainable Brownian paths $\mathbf{B}_{[0,1]}$, then unbiasedness does not hold anymore. What is guaranteed, though, is that the ratio

$$\frac{\mathbb{E}\left[\mathbf{GR}_k(\mathbf{B}_{[0,1]}) \mid IV_{[0,T]}\right]}{\mathbb{E}\left[\mathbf{GR}_k(\mathbf{B}_{[0,1]})\right]} \longrightarrow 1$$

has a limiting value of 1 due to the established a.s. limit $1/2k^{-1/2}\mathbf{GR}_k \longrightarrow 1$ and both conditional and unconditional L_2-boundedness of $1/2k^{-1/2}\mathbf{GR}_k$ Thus, in the absence of jumps and drift, unbiasedness is achieved in the limit as k gets large

$$\mathbb{E}\left[e_k'\mathbf{GR}_k(\mathbf{p}_{[0,T]}) \mid IV_{[0,T]}\right] \longrightarrow \sqrt{IV_{[0,T]}} \text{ as } k \longrightarrow \infty \ ,$$

while in the particular case of no leverage effect (and no jumps) unbiasedness holds for any k:

$$\mathbb{E}\left[e_k'\mathbf{GR}_k(\mathbf{p}_{[0,T]}) \mid IV_{[0,T]}\right] = \sqrt{IV_{[0,T]}}$$

Either way, when jumps and possibly non-zero drift are present, their impact on the value of the product $e_k'\mathbf{GR}_k(\mathbf{p}_{[0,T]})$ vanishes only asymptotically as detailed in the proof of proposition 7 above. Therefore, only asymptotic unbiasedness holds in the presence of jumps or non-zero drift, regardless of whether the log-price process is subject to "leverage effect" or not.

Q.E.D. ■

References

Aït-Sahalia, Y., P.A. Mykland and L. Zhang (2005a), "How Often to Sample a Continuous-Time Process in the Presence of Market Microstructure Noise," Review of Financial Studies, 18, 351-416.

Aït-Sahalia, Y., P.A. Mykland and L. Zhang (2005b), "Ultra High Frequency Volatility Estimation with Dependent Microstructure Noise," Manuscript, Princeton University.

Alizadeh, S., M. Brandt, and F. X. Diebold (2002), "Range-Based Estimation of Stochastic Volatility Models," Journal of Finance 57, 1047-1091.

Andersen, T.G., L. Benzoni and J. Lund (2002), "An Empirical Investigation of Continuous-Time Equity Return Models," Journal of Finance, 57, 1239-1284.

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

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

Andersen, T.G., T. Bollerslev and F.X. Diebold (2004), "Parametric and Non-Parametric Volatility Measurement," in Handbook of Financial Econometrics (L.P Hansen and Y. Aït-Sahalia, eds.). Elsevier Science, New York, forthcoming.

Andersen, T.G., T. Bollerslev and F.X. Diebold (2005), "Roughing It Up: Including Jump Components in the Measurement, Modeling and Forecasting of Return Volatility," forthcoming, Review of Economics and Statistics.

Andersen, T.G., T. Bollerslev, F.X. Diebold and H. Ebens (2001), "The Distribution of Realized Stock Return Volatility," Journal of Financial Economics, 61, 43-76.

Andersen, T.G., T. Bollerslev, F.X. Diebold and P. Labys (2001), "The Distribution of Realized Exchange Rate Volatility," Journal of the American Statistical Association, 96, 42-55.

Andersen, T.G., T. Bollerslev, F.X. Diebold and P. Labys (2003), "Modeling and Forecasting Realized Volatility," Econometrica, 71, 579-625.

Andersen, T.G., T. Bollerslev and D. Dobrev (2006), "No Arbitrage Semi-Martingale Restrictions for Continuous-Time Volatility Models subject to Leverage Effects, Jumps, and i.i.d. Noise: Theory and Testable Distributional Implications," Journal of Econometrics, forthcoming.

Andersen, T.G., T. Bollerslev, P.H. Frederiksen and M.Ø. Nielsen (2006), "Jumps in Common Stock Returns," Manuscript, Northwestern University, Duke University, and Cornell University.

Andersen, T. G., T. Bollerslev and X. Huang (2006), "A Semiparametric Framework for Modelling and Forecasting Jumps and Volatility in Speculative Prices," Manuscript, Duke University.

Ané, T. And H. Geman (2000), "Order Flow, Transaction Clock, and Normality of Asset Returns," Journal of Finance, 55, 2259-2284.

Back, K. (1991), "Asset Prices for General Processes," Journal of Mathematical Economics, 20, 317-395.

Bandi F.M. and J.R. Russell (2006), "Separating Microstructure Noise from Volatility," Journal of Financial Economics, 79 (3), 655-92.

Barndorff-Nielsen, O.E., P.R. Hansen, A. Lunde and N. Shephard (2005b), "Regular and Modified Kernel-Based Estimators of Integrated Variance: The Case with Independent Noise," Manuscript, Stanford University.

Barndorff-Nielsen, O.E., P.R. Hansen, A. Lunde and N. Shephard (2006), "Designing Realised Kernels to Measure the Ex-post Variation of Equity Prices in the Presence of Noise," Manuscript, Oxford Financial Research Centre.

Barndorff-Nielsen, O.E. and N. Shephard (2002a), "Econometric Analysis of Realised Volatility and its use in Estimating Stochastic Volatility Models," Journal of the Royal Statistical Society, 64, 253-280.

Barndorff-Nielsen, O.E. and N. Shephard (2002b), "Estimating Quadratic Variation Using Realized Variance," Journal of Applied Econometrics, 17, 457-478.

Barndorff-Nielsen, O.E. and N. Shephard (2004), "Power and Bipower Variation with Stochastic Volatility and Jumps," Journal of Financial Econometrics, 2, 1-37.

Barndorff-Nielsen, O.E. and N. Shephard (2006), "Econometrics of Testing for Jumps in Financial Economics Using Bipower Variation," Journal of Financial Econometrics, 4, 1-30.

Bollerslev, T., J. Litvinova and G. Tauchen (2005), "Leverage and Volatility Feedback Effects in High-Frequency Data," Manuscript, Duke University.

Brandt, M.W. and F.X. Diebold (2006), "A No-Arbitrage Approach to Range-Based Estimation of Return Covariances and Correlations," Journal of Business, 79, 61-74.

Carr P. and L. Wu (2003), "What Type of Process Underlies Options? A Simple Robust Test," Journal of Finance, 58 (6), 2581–2610.

Clark, P.K. (1973), "A Subordinate Stochastic Process Model with Finite Variance for Speculative Prices," Econometrica, 41, 135-155.

Christensen, K. and M. Podolskij (2006a), "Realized Range-Based Estimation of Integrated Variance," Journal of Econometrics, forthcoming.

Christensen, K. and M. Podolskij (2006b), "Range-Based Estimation of Quadratic Variation," Manuscript, Aarhus School of Business.

Dambis, K.E. (1965), "On the Decomposition of Continuous Submartingales," Theory of Probability Applications, 10, 401-410.

Dubins, L.E. and G. Schwartz (1965), "On Continuous Martingales," Proceedings of the National Academy of Sciences, 53, 913-916.

Dijk, D.V. and Martens, M. (2005), "Measuring Volatility with the Realized Range," Manuscript, Erasmus University, Rotterdam.

Dobrev, D. (1998), "Extreme Cumulative Returns From The Free Market Exchange of Assets," Master's Thesis, Sofia University.

Eraker, B. (2004), "Do Stock Prices and Volatility Jump? Reconciling Evidence from Spot and Option Prices," Journal of Finance, 59, 1367–1403.

Eraker, B., M. Johannes and N. Polson (2003), "The Impact of Jumps in Equity Index Volatility and Returns," Journal of Finance, 53, 1269-1300.

Feller, W. (1971), "An Introduction to Probability Theory and Its Applications," Vol. 2, 3rd ed. Wiley, New York.

Gallant, A.R., C.T. Hsu and G. Tauchen (1999), "Using Daily Range Data to Calibrate Volatility Diffusions and Extract the Forward Integrated Variance," Review of Economics and Statistics, 81, 617-631.

Garman, M.B. and M.J. Klass (1980), "On the Estimation of Price Volatility from Historical Data," Journal of Business, 53, 67-78.

Gatev E., W.N. Goetzmann and K.G. Rouwenhorst (2006), "Pairs Trading: Performance of a Relative-Value Arbitrage Rule," Review of Financial Studies 2006, 19 (3), 797-827.

Ghysels, E., A.C. Harvey and E. Renault (1996), "Stochastic Volatility," in Handbook of Statistics, Volume 14 (G.S. Maddala and C.R. Rao, eds.). North Holland, Amsterdam.

Hasbrouck, J. (1999), "The Dynamics of Discrete Bid and Ask Quotes," Journal of Finance, 54, 2109-2142.

Hansen, P.R. and A. Lunde (2006), "Realized Variance and Market Microstructure Noise," Journal of Business and Economic Statistics, 24, 127-218.

Huang, X. and G. Tauchen (2005), "The Relative Contribution of Jumps to Total Price Variance," Journal of Financial Econometrics, 3, 456-499.

Karatzas, I. and S.E. Shreve (1991), "Brownian Motion and Stochastic Calculus," 2nd ed. Springer Verlag, New York.

Kunitomo, N. (1992), "Improving the Parkinson Method of Estimating Security Price Volatilities," Journal of Business 65 (2), 295-302.

Li, K. and D. Weinbaum (2000), "The Empirical Performance of Alternative Extreme Value Volatility Estimators," Manuscript, New York University.

Magdon-Ismail M. and A.F. Atiya (2004), "Maximum Drawdown," Risk Magazine, 17 (10), 99-102.

Maheu, J.M. and McCurdy T.H. (2004), "News Arrival, Jump Dynamics and Volatility Components for Individual Stock Returns," Journal of Finance, 59, 755–793.

Meddahi, N. (2002), "A Theoretical Comparison Between Integrated and Realized Volatility," Journal of Applied Econometrics, 17, 479-508.

Merton, R.C. (1980), "On Estimating the Expected Return on the Market: An Exploratory Investigation," Journal of Financial Economics, 8, 323-361.

Neveu, J. and J. Pitman (1989), "Renewal Property of The Extrema and The Tree Property of a One-Dimensional Brownian motion,"Séminaire de Probabilités XXIII. Lecture Notes in Math., 1372, 239–247, Springer, New York.

Pan, J. (2002), "The Jump-Risk Premia Implicit in Options: Evidence from an Integrated Time-Series Study,"Journal of Financial Economics, 63, 3-50.

Parkinson, M. (1980), "The Extreme Value Method for Estimating the Variance of the Rate of Return,"Journal of Business, 53 (1), 61-65.

Revuz, D. and M. Yor (1998), "Continuous Martingales and Brownian Motion,"3rd ed. Springer, Berlin.

Rogers, L. C. G. (1998), "Volatility Estimation with Price Quanta,"Mathematical Finance, 8 (3), 277-290.

Rogers, L. C. G. and S. E. Satchell (1991), "Estimating variance from high, low and closing prices," Journal of Business, 1 (4), 504-512.

Tsirelson, B. (2006), "Brownian Local Minima, Random Dense Countable Sets and Random Equivalence Classes," Electronic Journal of Probability, 11 (7), 162-198

Vecer, J. (2006), "Maximum Drawdown and Directional Trading," Risk, 19 (12), 88-92

Zhang, L., Y. Aït-Sahalia and P.A. Mykland (2005), "A Tale of Two Time Scales: Determining Integrated Volatility with Noisy High-Frequency Data," Journal of the American Statistical Association 100, 1394-1411.

Table 1 Daily HAR-RV-CJ Volatility Forecast Regressions for 10 NYSE and 10 NASDAQ Stocks

$$\sqrt{RV_{t,t+1}} = \beta_0 + \beta_{CD}\sqrt{C_t} + \beta_{CW}\sqrt{C_{t-5,t}} + \beta_{CM}\sqrt{C_{t-22,t}} + \beta_{JD}\sqrt{J_t} + \beta_{JW}\sqrt{J_{t-5,t}} + \beta_{JM}\sqrt{J_{t-22,t}} + \varepsilon_{t,t+h}$$

	C	CAT	EK	GE	GM	HON	IBM	MCD	MO	X	ADSK	ADTN	CBRL	FITB	LRCX	HYSL	SIAL	SNPS	SPLS	TECD
β_0	0.029	0.177	0.243	0.069	0.140	0.096	0.058	0.139	0.106	0.467	0.222	0.576	0.121	0.195	0.141	0.183	0.118	0.301	0.121	0.120
	(0.038)	(0.048)	(0.063)	(0.031)	(0.044)	(0.044)	(0.030)	(0.048)	(0.064)	(0.107)	(0.084)	(0.157)	(0.063)	(0.044)	(0.067)	(0.099)	(0.046)	(0.088)	(0.064)	(0.072)
β_{CD}	0.622	0.299	0.538	0.353	0.329	0.322	0.295	0.405	0.307	0.410	0.365	0.265	0.256	0.380	0.327	0.308	0.434	0.286	0.263	0.291
	(0.134)	(0.042)	(0.072)	(0.066)	(0.057)	(0.060)	(0.061)	(0.077)	(0.079)	(0.061)	(0.051)	(0.050)	(0.071)	(0.078)	(0.041)	(0.075)	(0.057)	(0.050)	(0.050)	(0.044)
β_{CW}	0.265	0.486	0.140	0.373	0.474	0.565	0.592	0.251	0.455	0.288	0.300	0.229	0.189	0.421	0.350	0.371	0.194	0.204	0.415	0.259
	(0.100)	(0.095)	(0.072)	(0.134)	(0.093)	(0.094)	(0.105)	(0.093)	(0.166)	(0.071)	(0.075)	(0.089)	(0.081)	(0.115)	(0.085)	(0.074)	(0.082)	(0.087)	(0.105)	(0.075)
β_{CM}	0.121	0.104	0.226	0.204	0.125	0.086	0.066	0.297	0.208	0.080	0.348	0.247	0.515	0.080	0.277	0.251	0.319	0.405	0.226	0.446
	(0.071)	(0.059)	(0.077)	(0.085)	(0.064)	(0.092)	(0.050)	(0.073)	(0.104)	(0.075)	(0.094)	(0.105)	(0.089)	(0.073)	(0.067)	(0.092)	(0.074)	(0.091)	(0.076)	(0.078)
β_{JD}	0.007	0.063	0.023	-0.036	0.022	0.027	0.017	-0.017	0.063	0.007	-0.031	0.044	0.036	0.024	0.055	0.001	0.026	0.039	-0.024	0.056
	(0.121)	(0.054)	(0.029)	(0.049)	(0.032)	(0.051)	(0.048)	(0.029)	(0.072)	(0.031)	(0.040)	(0.035)	(0.033)	(0.044)	(0.039)	(0.038)	(0.039)	(0.033)	(0.034)	(0.041)
β_{JW}	-0.059	0.041	0.094	0.084	-0.080	0.030	0.056	-0.050	0.026	0.042	-0.012	0.010	0.122	0.057	0.060	-0.004	0.008	0.025	0.050	0.041
	(0.091)	(0.057)	(0.054)	(0.072)	(0.050)	(0.055)	(0.073)	(0.048)	(0.060)	(0.054)	(0.051)	(0.058)	(0.064)	(0.057)	(0.046)	(0.074)	(0.047)	(0.054)	(0.063)	(0.060)
β_{JM}	-0.104	-0.031	-0.084	0.089	0.126	-0.036	-0.005	-0.031	-0.082	0.101	-0.166	0.126	-0.097	-0.024	-0.005	0.112	0.026	-0.023	0.189	-0.076
	(0.072)	(0.073)	(0.069)	(0.085)	(0.100)	(0.056)	(0.073)	(0.088)	(0.049)	(0.068)	(0.099)	(0.094)	(0.102)	(0.072)	(0.070)	(0.101)	(0.078)	(0.092)	(0.118)	(0.072)
$R^2_{HAR\text{-}RV\text{-}CJ[GR]}$	0.799	0.570	0.454	0.724	0.567	0.617	0.758	0.516	0.373	0.334	0.424	0.211	0.402	0.481	0.632	0.430	0.528	0.364	0.586	0.447
$R^2_{HAR\text{-}RV\text{-}CJ[BV]}$	0.787	0.569	0.430	0.720	0.550	0.601	0.756	0.492	0.358	0.303	0.394	0.214	0.392	0.473	0.628	0.420	0.513	0.359	0.580	0.438

Key: The table reports OLS estimates for daily HAR-RV-CJ volatility forecast regressions with continuous and jump parts obtained from GR. All of the realized volatility measures and its continuous and jump parts are constructed from tick-by-tick TAQ quote data from July 2001 through December 2005, for a total of 1131 trading days. The weekly and monthly measures are the scaled sums of the corresponding daily measures. The standard errors reported in parentheses are based on a Newey-West correction allowing for serial correlation up to order 5. The last two rows labeled $R^2_{HAR\text{-}RV\text{-}CJ[GR]}$ and $R^2_{HAR\text{-}RV\text{-}CJ[BV]}$ are for the HAR-RV-CJ with continuous and jumps parts obtained from GR and BV respectively. The jump detection method using GR is based on this paper and Andersen, Bollerslev, and Dobrev (2006), while the method using BV is based on Andersen, Bollerslev, and Diebold (2006).

Table 2 Weekly HAR-RV-CJ Volatility Forecast Regressions for 10 NYSE and 10 NASDAQ Stocks

$$\sqrt{RV_{t,t+5}} = \beta_0 + \beta_{CD}\sqrt{C_t} + \beta_{CW}\sqrt{C_{t-5,t}} + \beta_{CM}\sqrt{C_{t-22,t}} + \beta_{JD}\sqrt{J_t} + \beta_{JW}\sqrt{J_{t-5,t}} + \beta_{JM}\sqrt{J_{t-22,t}} + \varepsilon_{t,t+h}$$

	C	CAT	EK	GE	GM	HON	IBM	MCD	MO	X	ADSK	ADTN	CBRL	FITB	LRCX	HYSL	SIAL	SNPS	SPLS	TECD
β_0	0.072	0.267	0.443	0.097	0.239	0.192	0.104	0.213	0.185	0.748	0.369	0.934	0.158	0.317	0.264	0.258	0.176	0.483	0.185	0.244
	(0.052)	(0.064)	(0.102)	(0.041)	(0.064)	(0.061)	(0.044)	(0.062)	(0.105)	(0.148)	(0.125)	(0.201)	(0.083)	(0.052)	(0.091)	(0.122)	(0.064)	(0.118)	(0.076)	(0.103)
β_{CD}	0.478	0.201	0.279	0.299	0.277	0.247	0.246	0.240	0.227	0.254	0.230	0.147	0.119	0.254	0.188	0.167	0.176	0.168	0.162	0.155
	(0.097)	(0.041)	(0.055)	(0.056)	(0.039)	(0.073)	(0.042)	(0.060)	(0.056)	(0.034)	(0.031)	(0.034)	(0.042)	(0.046)	(0.028)	(0.044)	(0.034)	(0.035)	(0.033)	(0.032)
β_{CW}	0.212	0.517	0.221	0.348	0.393	0.568	0.621	0.220	0.441	0.247	0.282	0.199	0.149	0.512	0.430	0.389	0.357	0.195	0.453	0.267
	(0.152)	(0.096)	(0.088)	(0.154)	(0.107)	(0.122)	(0.111)	(0.095)	(0.184)	(0.078)	(0.084)	(0.102)	(0.094)	(0.098)	(0.098)	(0.128)	(0.118)	(0.091)	(0.125)	(0.091)
β_{CM}	0.306	0.127	0.303	0.268	0.211	0.137	0.056	0.478	0.292	0.151	0.469	0.265	0.720	0.051	0.313	0.364	0.381	0.486	0.251	0.566
	(0.094)	(0.068)	(0.106)	(0.106)	(0.097)	(0.103)	(0.072)	(0.092)	(0.143)	(0.108)	(0.131)	(0.128)	(0.114)	(0.083)	(0.095)	(0.128)	(0.120)	(0.115)	(0.103)	(0.109)
β_{JD}	-0.092	0.022	0.027	-0.023	-0.008	0.006	0.066	-0.025	0.010	0.018	-0.006	0.006	0.032	0.014	0.032	0.023	0.001	0.037	-0.010	0.020
	(0.125)	(0.030)	(0.026)	(0.043)	(0.024)	(0.045)	(0.036)	(0.022)	(0.028)	(0.022)	(0.022)	(0.019)	(0.019)	(0.033)	(0.020)	(0.022)	(0.027)	(0.018)	(0.019)	(0.019)
β_{JW}	0.205	0.098	0.064	0.125	-0.044	0.094	-0.051	-0.019	0.053	0.021	-0.039	0.052	0.139	0.052	0.053	-0.101	0.056	-0.034	0.073	0.089
	(0.111)	(0.090)	(0.074)	(0.093)	(0.063)	(0.075)	(0.095)	(0.063)	(0.074)	(0.060)	(0.056)	(0.057)	(0.080)	(0.094)	(0.051)	(0.088)	(0.062)	(0.057)	(0.057)	(0.055)
β_{JM}	-0.290	-0.080	-0.061	0.119	0.125	-0.133	0.062	-0.107	-0.125	0.138	-0.195	0.133	-0.170	-0.043	-0.028	0.195	0.033	-0.018	0.235	-0.168
	(0.137)	(0.102)	(0.115)	(0.122)	(0.137)	(0.077)	(0.110)	(0.105)	(0.069)	(0.095)	(0.132)	(0.114)	(0.125)	(0.090)	(0.105)	(0.169)	(0.111)	(0.115)	(0.138)	(0.098)
$R^2_{HAR\text{-}RV\text{-}CJ[GR]}$	0.786	0.662	0.432	0.785	0.638	0.691	0.818	0.606	0.455	0.344	0.512	0.254	0.565	0.560	0.731	0.573	0.637	0.472	0.736	0.567
$R^2_{HAR\text{-}RV\text{-}CJ[BV]}$	0.775	0.661	0.416	0.784	0.618	0.678	0.814	0.593	0.435	0.327	0.480	0.290	0.557	0.546	0.727	0.553	0.633	0.466	0.740	0.541

Key: The table reports OLS estimates for weekly HAR-RV-CJ volatility forecast regressions with continuous and jump parts obtained from GR. All of the realized volatility measures and its continuous and jump parts are constructed from tick-by-tick TAQ quote data from July 2001 through December 2005, for a total of 1131 trading days. The weekly and monthly measures are the scaled sums of the corresponding daily measures. The standard errors reported in parentheses are based on a Newey-West correction allowing for serial correlation up to order 10. The last two rows labeled $R^2_{HAR\text{-}RV\text{-}CJ[GR]}$ and $R^2_{HAR\text{-}RV\text{-}CJ[BV]}$ are for the HAR-RV-CJ with continuous and jumps parts obtained from GR and BV respectively. The jump detection method using GR is based on this paper and Andersen, Bollerslev, and Dobrev (2006), while the method using BV is based on Andersen, Bollerslev, and Diebold (2006).

Table 3 Monthly HAR-RV-CJ Volatility Forecast Regressions for 10 NYSE and 10 NASDAQ Stocks

$$\sqrt{RV_{t,t+22}} = \beta_0 + \beta_{CD}\sqrt{C_t} + \beta_{CW}\sqrt{C_{t-5,t}} + \beta_{CM}\sqrt{C_{t-22,t}} + \beta_{JD}\sqrt{J_t} + \beta_{JW}\sqrt{J_{t-5,t}} + \beta_{JM}\sqrt{J_{t-22,t}} + \varepsilon_{t,t+h}$$

	C	CAT	EK	GE	GM	HON	IBM	MCD	MO	X	ADSK	ADTN	CBRL	FITB	LRCX	HYSL	SIAL	SNPS	SPLS	TECD
β_0	0.169	0.465	0.738	0.192	0.478	0.443	0.224	0.338	0.416	1.303	0.657	1.443	0.226	0.628	0.523	0.426	0.323	0.774	0.371	0.410
	(0.083)	(0.077)	(0.134)	(0.069)	(0.113)	(0.111)	(0.072)	(0.086)	(0.117)	(0.208)	(0.171)	(0.223)	(0.108)	(0.102)	(0.159)	(0.176)	(0.092)	(0.187)	(0.111)	(0.140)
β_{CD}	0.289	0.150	0.145	0.213	0.172	0.176	0.214	0.126	0.154	0.106	0.138	0.075	0.052	0.186	0.136	0.074	0.119	0.104	0.112	0.105
	(0.032)	(0.027)	(0.029)	(0.035)	(0.031)	(0.034)	(0.037)	(0.032)	(0.036)	(0.020)	(0.019)	(0.020)	(0.012)	(0.035)	(0.024)	(0.019)	(0.023)	(0.021)	(0.027)	(0.023)
β_{CW}	0.277	0.321	0.191	0.286	0.316	0.382	0.434	0.164	0.433	0.193	0.159	0.116	0.126	0.326	0.352	0.157	0.285	0.212	0.309	0.145
	(0.125)	(0.112)	(0.077)	(0.124)	(0.113)	(0.097)	(0.104)	(0.057)	(0.157)	(0.070)	(0.073)	(0.064)	(0.072)	(0.083)	(0.089)	(0.103)	(0.084)	(0.051)	(0.079)	(0.090)
β_{CM}	0.390	0.233	0.298	0.348	0.216	0.282	0.152	0.593	0.225	0.103	0.604	0.306	0.855	0.150	0.369	0.553	0.434	0.373	0.290	0.707
	(0.152)	(0.103)	(0.107)	(0.106)	(0.163)	(0.088)	(0.120)	(0.067)	(0.155)	(0.144)	(0.179)	(0.121)	(0.150)	(0.120)	(0.103)	(0.219)	(0.132)	(0.118)	(0.107)	(0.130)
β_{JD}	-0.014	0.024	-0.004	-0.006	0.010	0.024	0.040	-0.003	0.007	0.006	-0.009	0.019	0.026	0.003	0.024	0.002	0.017	0.012	-0.001	0.007
	(0.065)	(0.018)	(0.010)	(0.023)	(0.018)	(0.022)	(0.023)	(0.013)	(0.015)	(0.013)	(0.011)	(0.015)	(0.011)	(0.028)	(0.012)	(0.010)	(0.012)	(0.009)	(0.008)	(0.010)
β_{JW}	0.003	0.029	-0.041	0.135	-0.040	0.001	-0.110	-0.068	-0.044	0.012	-0.055	0.116	0.066	0.069	0.023	-0.050	0.053	-0.022	0.047	0.023
	(0.116)	(0.069)	(0.048)	(0.099)	(0.101)	(0.072)	(0.130)	(0.042)	(0.045)	(0.043)	(0.063)	(0.056)	(0.061)	(0.072)	(0.048)	(0.054)	(0.042)	(0.043)	(0.052)	(0.049)
β_{JM}	-0.349	-0.005	0.053	0.134	0.198	-0.158	0.297	-0.147	-0.097	0.137	-0.297	-0.126	-0.288	-0.265	-0.079	0.301	-0.013	0.079	0.444	-0.218
	(0.398)	(0.144)	(0.181)	(0.178)	(0.253)	(0.120)	(0.285)	(0.135)	(0.113)	(0.130)	(0.228)	(0.129)	(0.152)	(0.153)	(0.179)	(0.173)	(0.225)	(0.155)	(0.165)	(0.126)
$R^2_{HAR\text{-}RV\text{-}CJ[GR]}$	0.696	0.528	0.366	0.722	0.517	0.544	0.699	0.624	0.416	0.225	0.500	0.212	0.654	0.417	0.677	0.637	0.614	0.480	0.686	0.610
$R^2_{HAR\text{-}RV\text{-}CJ[BV]}$	0.687	0.527	0.350	0.721	0.523	0.526	0.696	0.612	0.380	0.221	0.449	0.312	0.612	0.388	0.672	0.634	0.620	0.476	0.702	0.552

Key: The table reports OLS estimates for monthly HAR-RV-CJ volatility forecast regressions with continuous and jump parts obtained from GR. All of the realized volatility measures and its continuous and jump parts are constructed from tick-by-tick TAQ quote data from July 2001 through December 2005, for a total of 1131 trading days. The weekly and monthly measures are the scaled sums of the corresponding daily measures. The standard errors reported in parentheses are based on a Newey-West correction allowing for serial correlation up to order 44. The last two rows labeled $R^2_{HAR\text{-}RV\text{-}CJ[GR]}$ and $R^2_{HAR\text{-}RV\text{-}CJ[BV]}$ are for the HAR-RV-CJ with continuous and jumps parts obtained from GR and BV respectively. The jump detection method using GR is based on this paper and Andersen, Bollerslev, and Dobrev (2006), while the method using BV is based on Andersen, Bollerslev, and Diebold (2006).

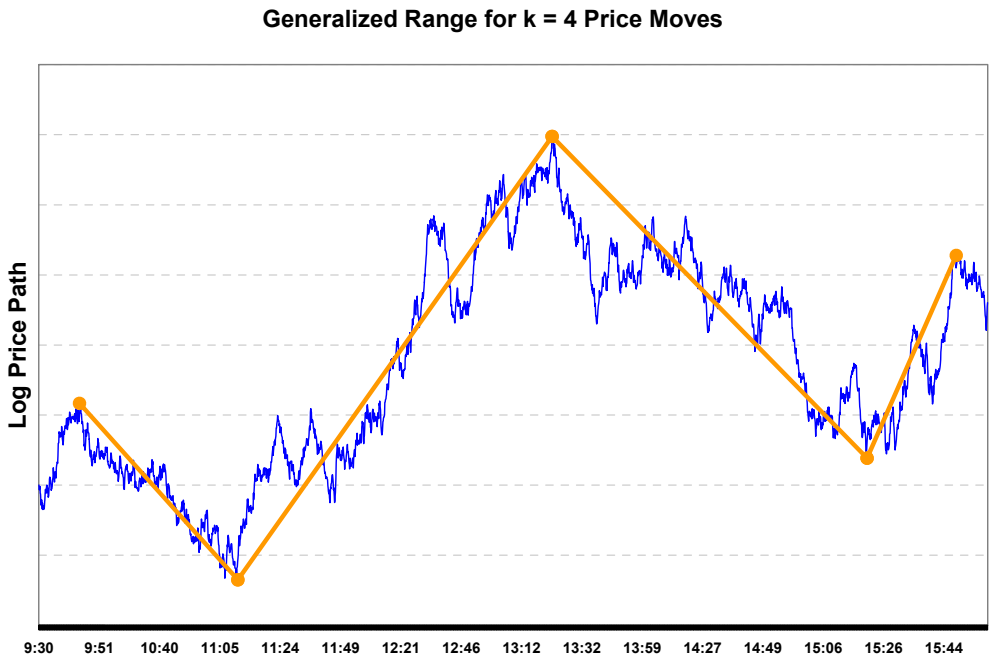


Figure 1. The figure illustrates the generalized range for k=4 price moves represented as line segments on a sample log-price path. The generalized range of any order k is defined as the maximized sum of k non-overlapping price moves on the observed path.

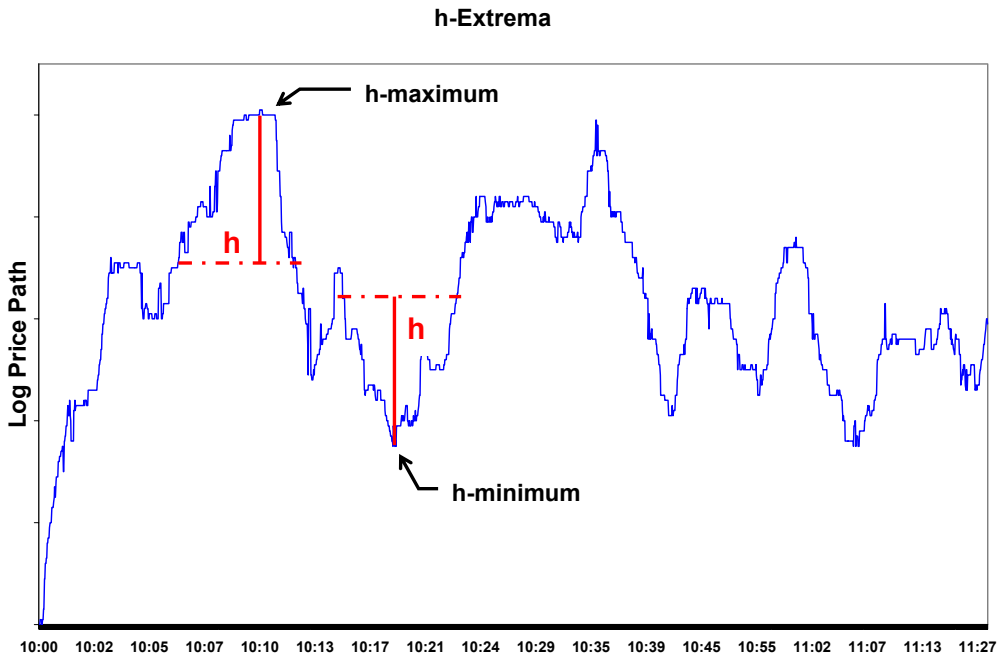


Figure 2. The figure illustrates an h-maximum followed by an h-minimum for an indicated threshold level h on a sample log-price path. A point is said to be an h-minimum (maximum) of a function if it is contained in an interval on which it is the global minimum (maximum) and the value taken by the function at both end points of the interval is above (below) this global minimum (maximum) at least by a threshold h.

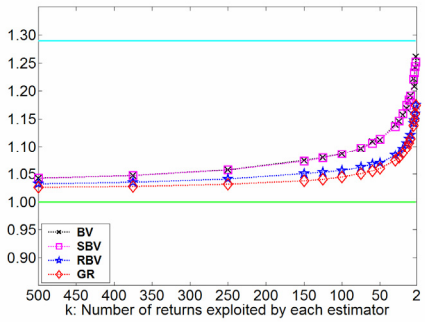
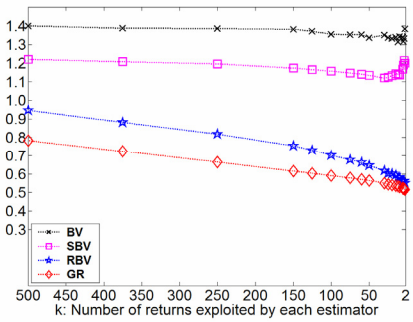
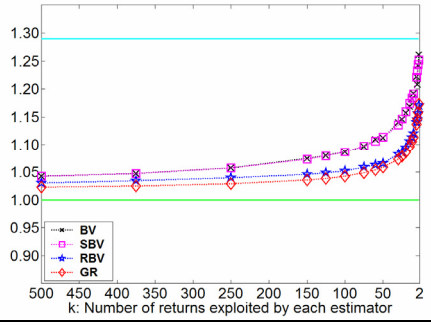
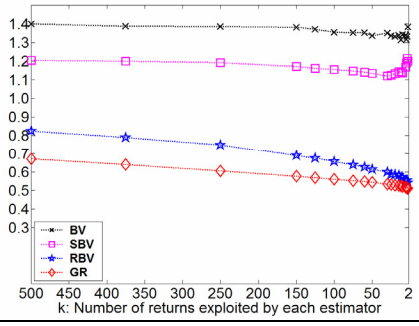
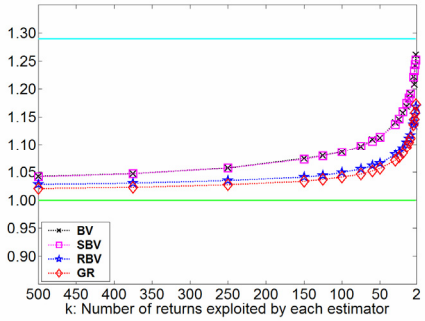
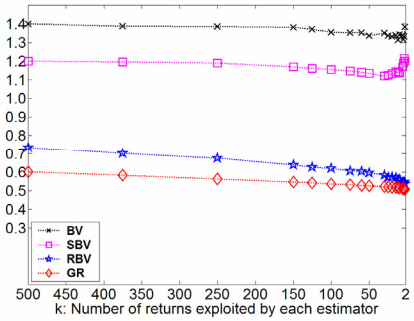
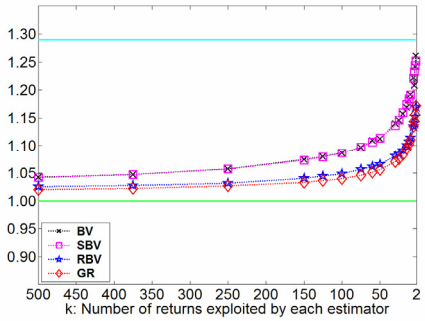
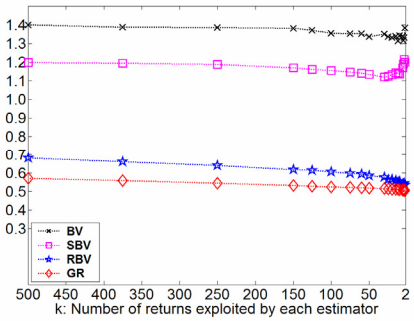
Sample Size	Mean Plots	Relative RMSE Plots
n = 1,500		
n = 3,000		
n = 7,500		
n =15,000		

Figure 3a. Mean and RMSE of the BV, SBV, RBV, and GR estimators of the diffusive part of total volatility as a function of the number of returns k = 1, 2, …, 500 exploited by each estimator for four different fixed sample sizes: n = 1,500; 3,000; 7,500; 15,000. The plots are based on 10,000 simulated daily jump-diffusion paths with roughly 20% jump contribution to total daily volatility and subject to leverage effect.

The common lower bound at level 1 for all plotted means is the true mean diffusive volatility, while the upper bound is the total volatility, including the jump contribution. The observed declining pattern of the upward bias of all diffusive volatility estimators as k gets large is in accordance with their asymptotic unbiasedness and can be interpreted as a "jump signature plot".

The RMSE of each estimator is expressed relative to the known asymptotic variance of the corresponding BV estimator in the absence of jumps having convergence rate $k^{1/2}$. For all plotted values of n and k the RMSE reduction GR vs. BV is up to 60%, the RMSE reduction GR vs. SBV is up to 50%, the RMSE reduction GR vs. RBV is up to 20%.

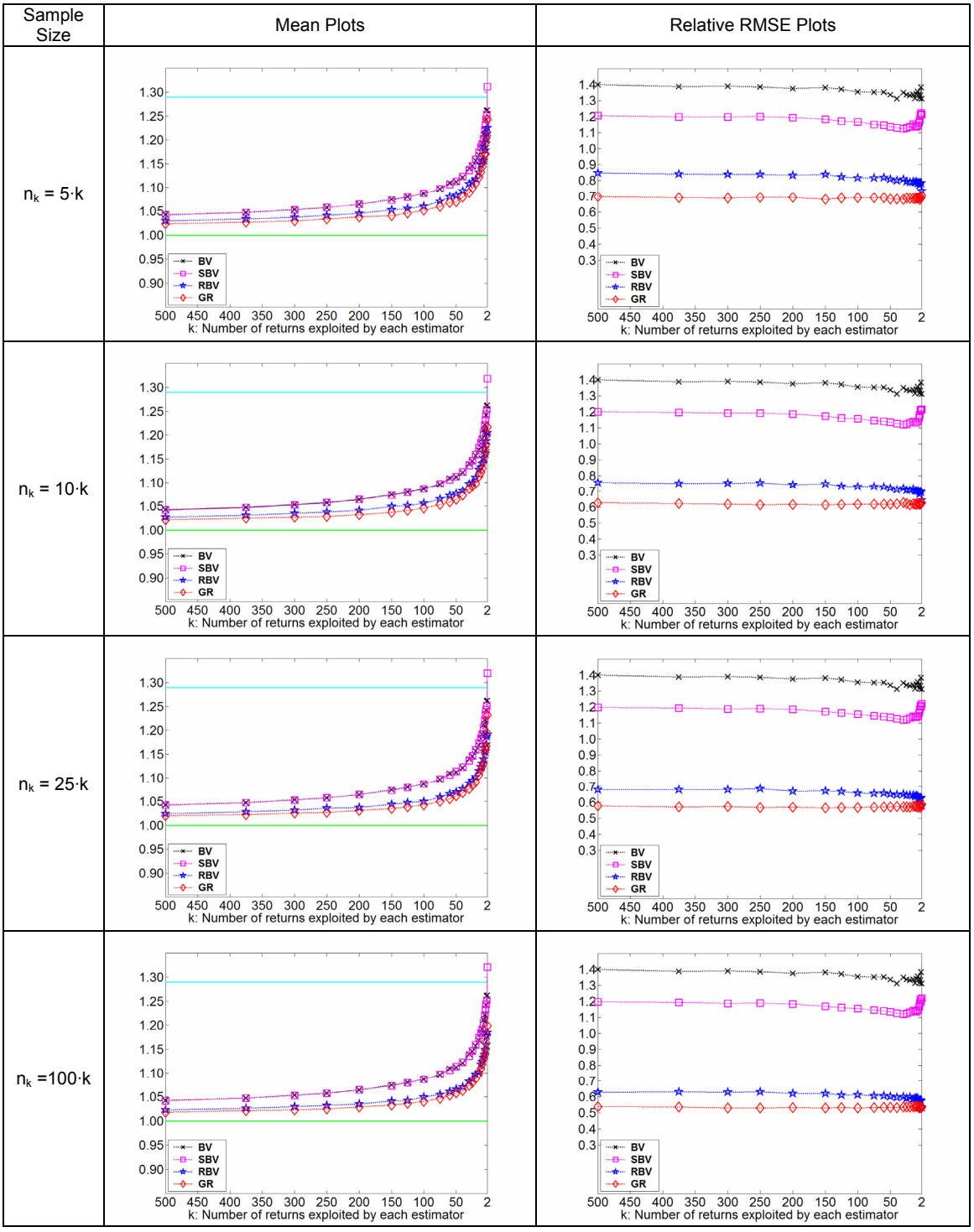


Figure 3b. Mean and RMSE of the BV, SBV, RBV, and GR estimators of the diffusive part of total volatility as a function of the number of returns k = 1, 2, …, 500 exploited by each estimator for four different sample sizes proportional to k: n_k = 5k; 10k; 25k; 100k. The plots are based on 10,000 simulated daily jump-diffusion paths with roughly 20% jump contribution to total daily volatility and subject to leverage effect.

The common lower bound at level 1 for all plotted means is the true mean diffusive volatility, while the upper bound is the total volatility, including the jump contribution. The observed declining pattern of the upward bias of all diffusive volatility estimators as k gets large is in accordance with their asymptotic unbiasedness and can be interpreted as a "jump signature plot".

The RMSE of each estimator is expressed relative to the known asymptotic variance of the corresponding BV estimator in the absence of jumps having convergence rate $k^{1/2}$. For all plotted values of n_k and k the RMSE reduction GR vs. BV is up to 60%, the RMSE reduction GR vs. SBV is up to 50%, and the RMSE reduction GR vs. RBV is up to 20%.

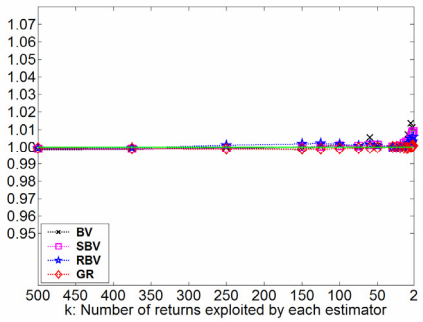
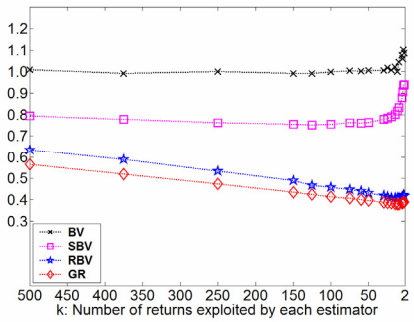
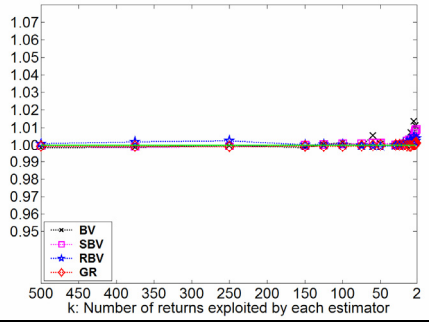
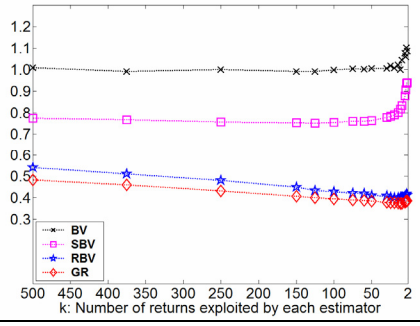
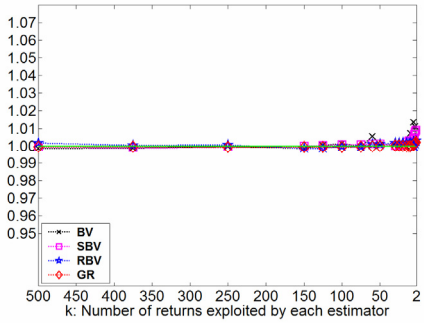
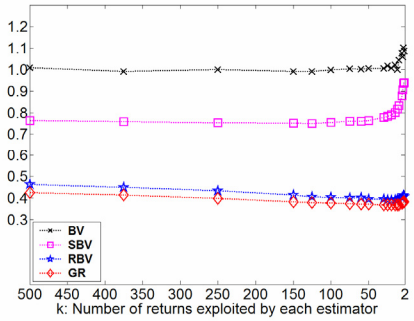
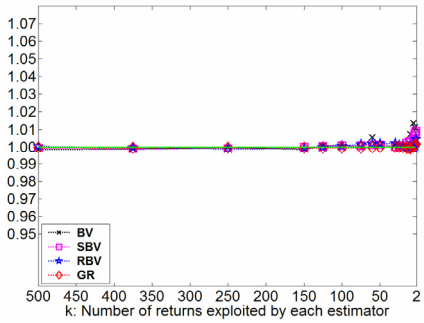
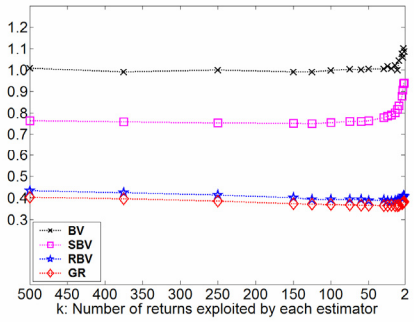
Sample Size	Mean Plots	Relative RMSE Plots
n = 1,500		
n = 3,000		
n = 7,500		
n =15,000		

Figure 4a. Mean and RMSE of the BV, SBV, RBV, and GR estimators of the diffusive part of total volatility as a function of the number of returns k = 1, 2, …, 500 exploited by each estimator for four different fixed sample sizes: n = 1,500; 3,000; 7,500; 15,000. The plots are based on 10,000 simulated daily jump-diffusion paths with 0% jump contribution to total daily volatility (i.e. in the absence of jumps) and subject to leverage effect.
The common mean level 1 for all plotted means is the true mean diffusive volatility, indicating that all estimators are equally unbiased when there are no jumps.
The RMSE of each estimator is expressed relative to the known asymptotic variance of the corresponding BV estimator in the absence of jumps having convergence rate $k^{1/2}$. For all plotted values of n and k the RMSE reduction GR vs. BV is up to 55%, the RMSE reduction GR vs. SBV is up to 45%, and the RMSE reduction GR vs. RBV is up to 15%.

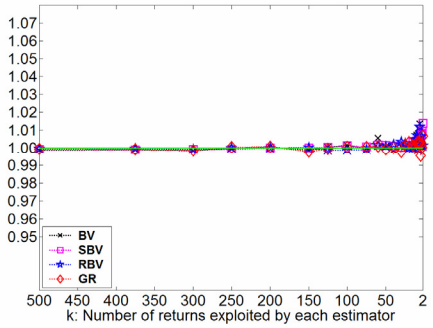
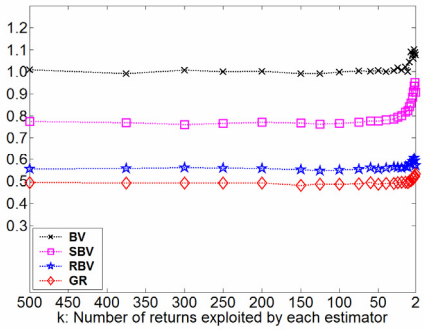
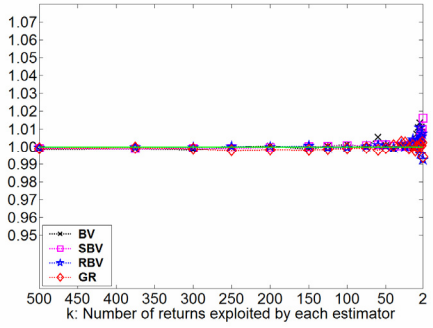
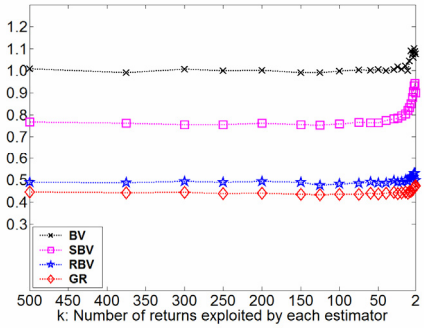
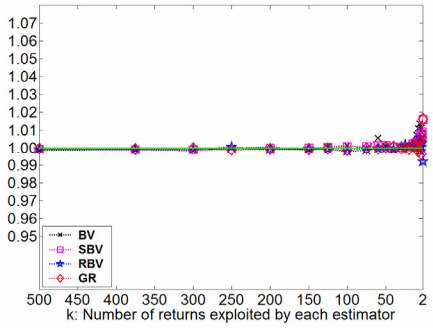
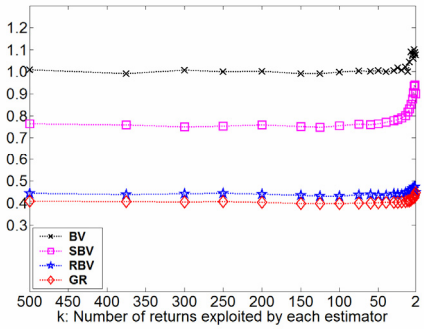
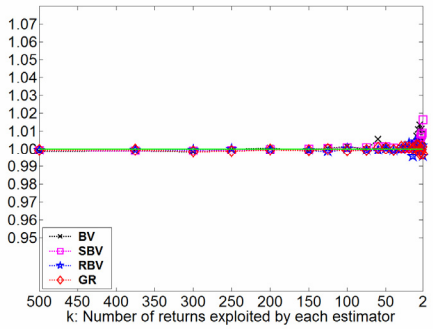
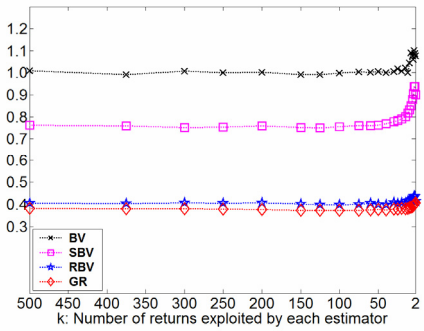
Sample Size	Mean Plots	Relative RMSE Plots
$n_k = 5 \cdot k$		
$n_k = 10 \cdot k$		
$n_k = 25 \cdot k$		
$n_k = 100 \cdot k$		

Figure 4b. Mean and RMSE of the BV, SBV, RBV, and GR estimators of the diffusive part of total volatility as a function of the number of returns k = 1, 2, …, 500 exploited by each estimator for four different sample sizes proportional to k: n_k = 5k; 10k; 25k; 100k. The plots are based on 10,000 simulated daily jump-diffusion paths with 0% jump contribution to total daily volatility (i.e. in the absence of jumps) and subject to leverage effect.
The common mean level 1 for all plotted means is the true mean diffusive volatility, indicating that all estimators are equally unbiased when there are no jumps.
The RMSE of each estimator is expressed relative to the known asymptotic variance of the corresponding BV estimator in the absence of jumps having convergence rate $k^{1/2}$. For all plotted values of n and k the RMSE reduction GR vs. BV is up to 55%, the RMSE reduction GR vs. SBV is up to 45%, and the RMSE reduction GR vs. RBV is up to 15%.

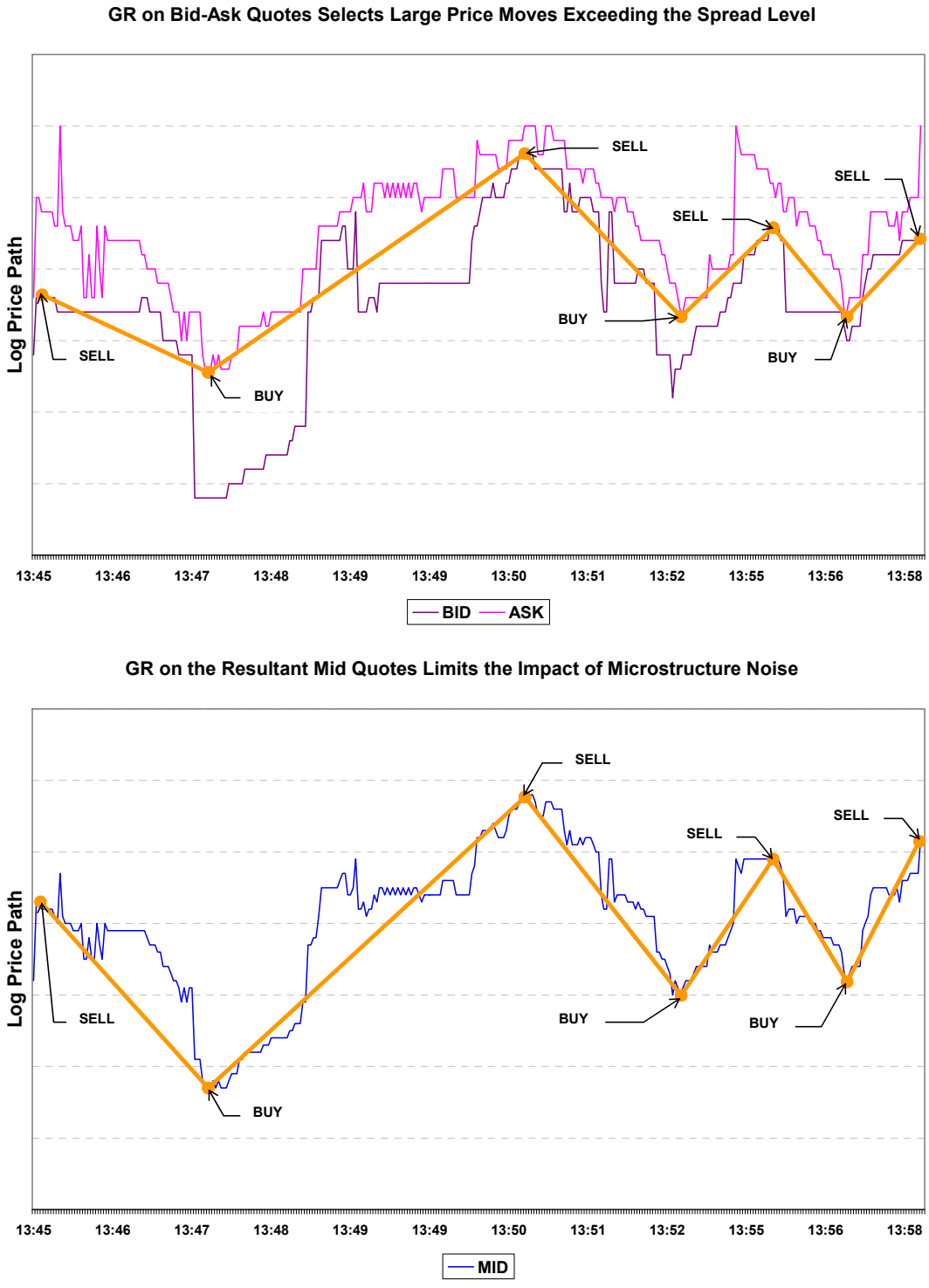


Figure 5. The figure illustrates how to calculate GR first on bid-ask stock quotes and then on the corresponding mid quotes in order to avoid distortion by microstructure noise. This two-step procedure makes an automatic selection of the maximum number of price moves in GR above the noise level, as determined by the spread. The plotted intraday quotes are for Helmerich & Payne Inc (NYSE:HP) from 13:45 to 14:00 on May 5, 2005, as recorded in TAQ.

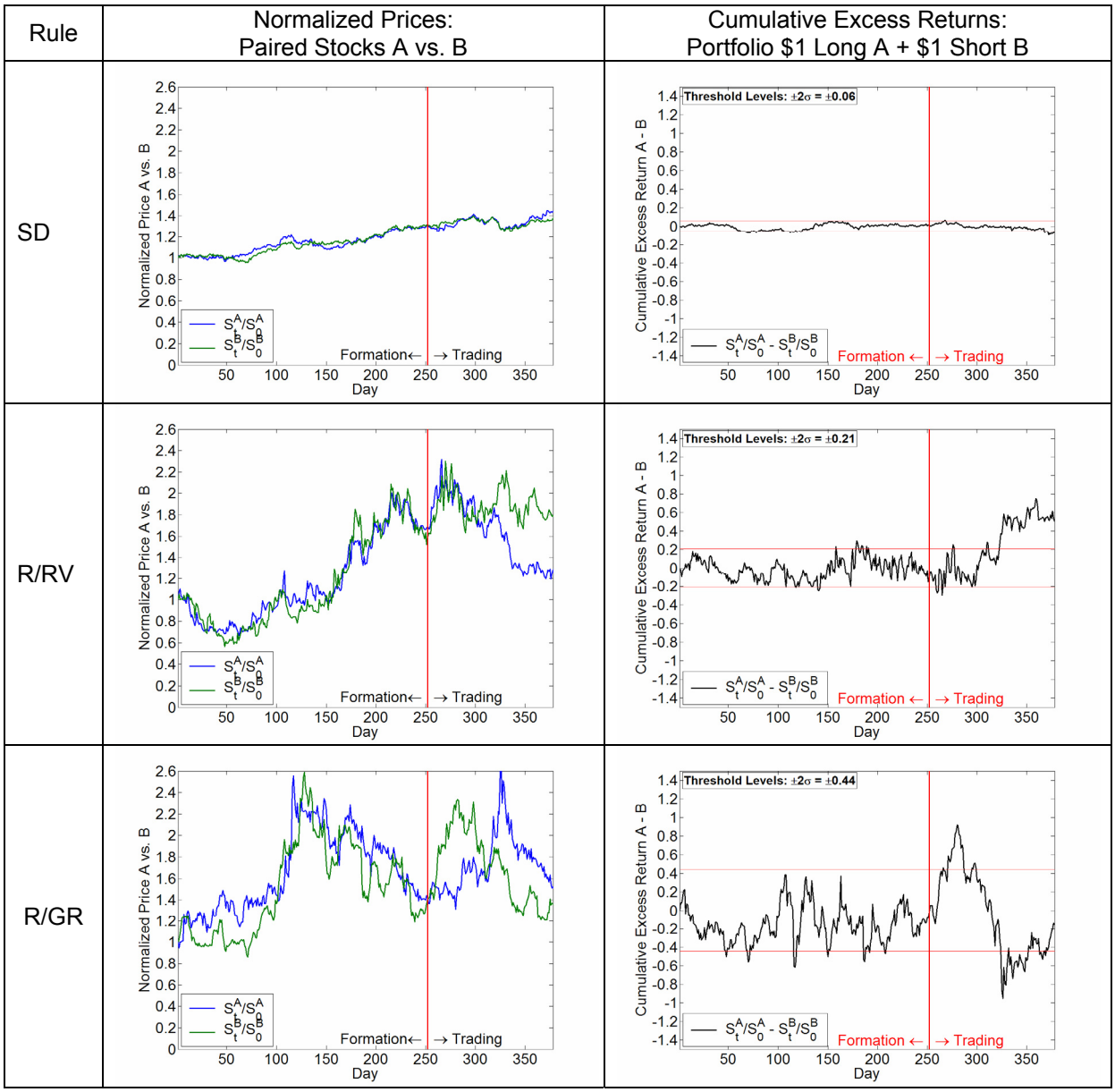


Figure 6. The figure plots the normalized price paths and corresponding cumulative excess returns of the long-short portfolio A – B for the twentieth selected stock pair {A, B} by three different rules: SD (minimum sum of squared deviations), RV (minimum ratio range to realized volatility), GR (minimum ratio range to generalized range of order six). Each plot shows the paths during a formation period of 252 days (year 2003) and the following trading period of 126 days (first half of 2004). The left column plots the price evolution of stock A vs. stock B normalized to start at $1 (i.e. cumulative returns). The right column plots the excess returns of the portfolio $1 long A and $1 short B. The red horizontal lines indicate the corresponding trading threshold level ±$2σ (two historical standard deviations, also reported in the upper left corner).

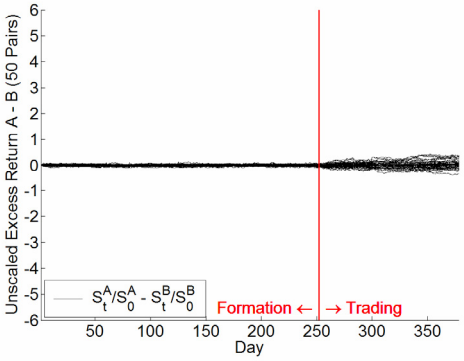
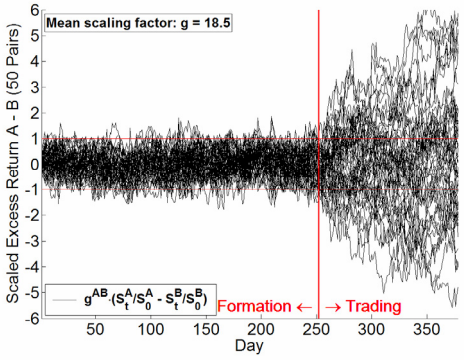
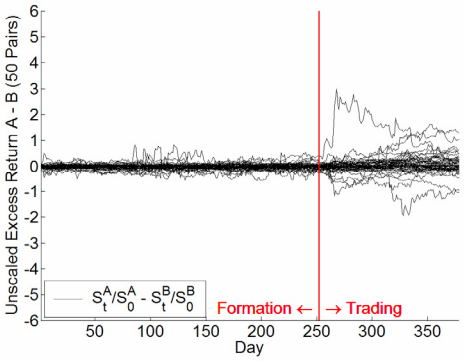
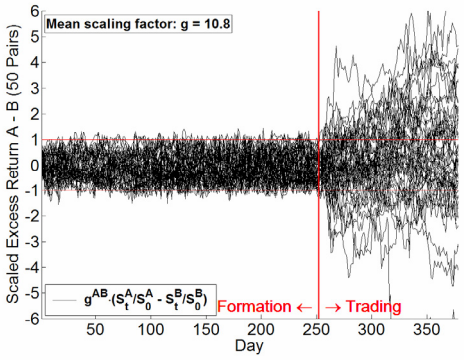
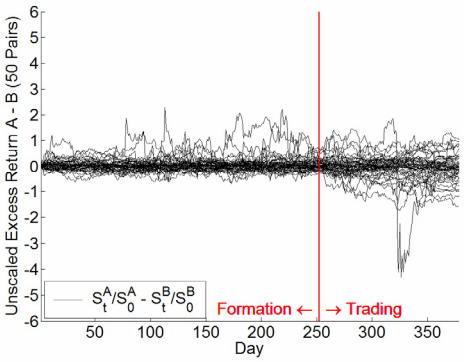
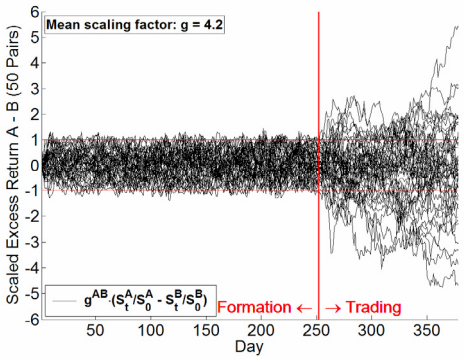
Rule	Unscaled Cumulative Excess Returns: 50 Portfolios $1 Long A + $1 Short B	Scaled Cumulative Excess Returns: 50 Portfolios $g^{A,B} Long A + $g^{A,B} Short B
SD		
R/RV		
R/GR		

Figure 7. The figure plots the excess return paths for the 50 best long-short pair portfolios A – B selected by three different rules: SD (minimum sum of squared deviations), RV (minimum ratio range to realized volatility), GR (minimum ratio range to generalized range). Each plot shows the paths during a formation period of 252 days (year 2003) and the following trading period of 126 days (first half of 2004). The left column plots the unscaled excess returns of each pair portfolio $1 long A and $1 short B. The right column plots the excess returns after scaling each portfolio to $g^{A,B} long A and $g^{A,B} short B, where the scaling factor $g^{A,B}$ for each pair is chosen to achieve a common trading threshold ±$1 (red horizontal lines), defined as two standard deviations during the formation period. The mean scaling factor g for the selected pairs is reported in the upper left corner.