

CREATES Research Paper 2007-19

Risk, Jumps, and Diversification

Tim Bollerslev, Tzuo Hann Law and George Tauchen



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



Aarhus School of Business
University of Aarhus

Handelsøjskolen
Aarhus Universitet

UNIVERSITY OF
COPENHAGEN



Risk, Jumps, and Diversification

Tim Bollerslev, Tzuo Hann Law and George Tauchen ^{*†}

Duke University

First Draft: March 1, 2007

May 29, 2007

Abstract

We test for price discontinuities, or jumps, in a panel of high-frequency intraday returns for forty large-cap stocks and an equiweighted index from these same stocks. Jumps are naturally classified into two types: common and idiosyncratic. Common jumps affect all stocks, albeit to varying degrees, while idiosyncratic jumps are stock-specific. Despite the fact that each of the stocks has a β of about unity with respect to the index, common jumps are virtually never detected in the individual stocks. This is truly puzzling, as an index can jump only if one or more of its components jump. To resolve this puzzle, we propose a new test for cojumps. Using this new test we find strong evidence for many modest-sized common jumps that simply pass through the standard jump detection statistic, while they appear highly significant in the cross section based on the new cojump identification scheme. Our results are further corroborated by a striking within-day pattern in the non-diversifiable cojumps.

^{*}Authors' contacts: bolle@econ.duke.edu; tzuohann.law@duke.edu; and george.tauchen@duke.edu.

[†]The work of Bollerslev was supported by a grant from the NSF to the NBER and CREATES funded by the Danish National Research Foundation. We thank members of the Duke University Econometrics and Finance Lunch Group and participants at the April 21-22, 2007 Stevanovich conference on "Volatility and High Frequency Data" and the May 18-19, 2007, "Financial Econometrics Conference" at Imperial College, London for helpful comments. We are especially grateful to Viktor Todorov and Xin Huang for many discussions and encouragement along the way.

1 Introduction

We examine the relationship between jumps in individual stocks and jumps in an aggregate market index. Several studies have recently presented strong nonparametric high-frequency data based empirical evidence in favor of jumps in financial asset prices, thus discrediting the classical continuous time paradigm with continuous sample price paths in favor of one with jump discontinuities.¹ The tests that we implement here further corroborate this evidence for the presence of jumps at both the individual stock and aggregate market level. Of particular interest, however, is the contrast between the outcome of tests for jumps conducted at the level of the individual stocks and the level of the index. This contrast presents an intriguing puzzle that we in turn resolve through the development of a new test for *cojumps*. The resolution provides key economic insights into the nature of stock price jumps and the role of diversification in mitigating the effects of jumps at the portfolio level.

Jumps are clearly of importance for asset allocation and risk management. A risk averse investor might be expected to shun investments with sharp unforeseeable movements. As an example of a jump, consider Figure 1, which depicts the price of Proctor and Gamble (PG) on August 3, 2004 sampled once every thirty seconds. There is a sharp discontinuity in the evolution of the prices a little after 11am, where PG gains about 30 cents over a period of just two minutes. Jumps like that are of great importance for standard arbitrage based arguments and derivatives pricing in particular, as the effect cannot readily be hedged by a portfolio of the underlying asset, cash, and other derivatives.²

Meanwhile, not all jumps are as clearly *ex post* identifiable as that shown in Figure 1, so that a formal statistical methodology for identifying jumps is needed. In the results reported on below we rely on the jump statistics developed in the seminal work by Barndorff-Nielsen and Shephard (2004) (BN-S). The BN-S theory provides a convenient nonparametric framework for measuring the relative contribution of jumps to total return variation and for classifying days on which jumps have or have not occurred. Even though this procedure has arguably emerged as the most popular high-frequency based jump detection scheme, it is important to keep in mind

¹Empirical studies and new statistical procedures include the work of Ait-Sahalia and Jacod (2006), Andersen et al. (2007), Barndorff-Nielsen and Shephard (2004), Fan and Wang (2006), Huang and Tauchen (2005), Jiang and Oomen (2006), Lee and Mykland (2007), and Mancini (2004).

²More formally, the effect of a jump on the price of a derivative is locally nonlinear and thereby cannot be neutralized by holding an *ex ante*-determined portfolio of other assets.

that the BN-S approach is an inferential procedure, and as such it sometimes missclassifies. It is in essence this failure of the test to detect certain jumps at the level of the individual stocks vis-a-vis the jumps that are detectable at the aggregate market level that creates an anomaly and thereby provides new insight into the functioning of the underlying markets.

Our empirical investigations are based on high-frequency intraday returns for a sample of forty large-cap U.S. stocks and the corresponding equiweighted index of these same stocks over the 2001–2005 sample period. Consistent with previous empirical results, we find strong evidence for the presence of jumps in each of the individual stock price series as well as the aggregate index. Standard computations also indicate about 15–25 large jumps for each of the individual stocks scattered randomly across the five-year sample, with jumps accounting for 12 percent of the total variation on average. In contrast, for the aggregate index there are only seven highly significant large jumps across the whole sample, and jumps as a whole account for just about 9 percent of the total return variation. Although these specific numbers obviously rely on our use of the popular BN-S procedure for identifying jumps, the same basic findings of more frequent and larger sized jumps for the individual stocks compared to the index is entirely consistent with the limited empirical evidence based on the univariate threshold type statistic recently reported in Lee and Mykland (2007).

The fact that the data reveal a less important role for jumps in the index than in each of its components is not all surprising. Basic portfolio theory implies that the idiosyncratic jumps should be diversified away in the aggregate portfolio. What is anomalous, however, is the apparent total lack of association between the significant jumps in each of the individual stocks and the jumps in the index. The basic jump detection procedure relies on a daily z -statistic, with large positive values of the statistic discrediting the null hypothesis of no jumps on that day. Oddly enough, the z -statistics for the individual stocks are essentially uncorrelated with the z -statistics for the index. This despite the fact that most of the individual stocks have β 's close to unity with respect to the index. Moreover, on the specific days for which the jump z -statistic for the index is statistically significant, indicating at least one jump, few if any of the z -statistics for the individual stocks are statistically significant, indicating no jumps. Of course, the equiweighted index can jump only if one or more of its components jumps.

The resolution of this apparent anomaly is the main focus of the paper. As we discuss below, the explanation lies in considering the magnitudes and correlation structures of the common

jumps, or what we term cojumps. To this end, we develop a new cojump test applicable to situations involving large panels of high frequency returns.³ Using this new test statistic we find strong evidence for many modest-sized common jumps that simply pass through the standard jump detection statistic, while they appear highly significant in the cross section based on the new cojump identification procedure.

As noted by Andersen et al. (2007), Barndorff-Nielsen and Shephard (2006), Eraker et al. (2003), and Huang and Tauchen (2005) among others, many, although not all, of the statistically significant jumps in aggregate stock indexes coincide with macroeconomic news announcements and other *ex-post* readily identifiable broad based economic news which similarly impact financial markets in a systematic fashion.⁴ While macroeconomic events obviously also affect individual firms, individual stock prices are also affected by sudden unexpected firm-specific information that can force an abrupt revaluation of the firms' stock.⁵ Our results suggest that firm-specific news events are indeed the dominant effect of the two in terms of their immediate price impact at the individual stock level, and only by properly considering the cross-section of returns do the non-diversifiable cojumps become visible in a formal statistical sense. These results are further corroborated by our findings of strong intraday patterns in the importance of jumps across the day, with the peak in the pattern for the aggregate index closely aligned with the time of the release of regularly scheduled news announcements.

The rest of the paper proceeds as follows. Section 2 describes the BN-S test procedure and the corresponding realized variation measures. Section 3 discusses the high-frequency data and sampling schemes underlying our empirical analysis, with some of the details relegated to a Data Appendix. The main empirical puzzle related to the apparent disconnect between the significant jumps in the individual stocks and the jumps in the aggregate market index is presented in Section 4. Section 5 details our resolution of the puzzle and the new test for cojumps. Section 6 concludes with a few final remarks and suggestions for future research.

³Alternative statistical procedures for identifying common jump arrivals in pairs of returns have recently been developed by Jacod and Todorov (2007)

⁴Specific examples include the monthly employment report, FED interest rate changes, oil prices, legislative alterations, and security concerns.

⁵Specific examples include lawsuits against a cigarette company, announcements of war for a defense company, and legislation of privacy issues that affect an Internet search engine.

2 Theoretical Framework

2.1 Assumptions and and Notation

We consider a scalar log-price process $p(t)$, evolving in continuous time as

$$dp(t) = \mu(t)dt + \sigma(t)dw(t) + d\mathcal{L}_J(t), \quad (1)$$

where $\mu(t)$ and $\sigma(t)$ refer to the drift and local volatility, respectively, $w(t)$ is a standard Brownian motion, and $\mathcal{L}_J(t)$ is a pure jump Lévy process.⁶ We adopt the timing convention that one time-unit corresponds to a trading day, so that $\{p(t-1+s)\}_{s \in [0,1]}$ represents the continuous log-price record over trading day t , where the integer values $t = 1, 2, 3, \dots$ coincide with the end of the day.

In practice, of course, the price process is only sampled at a finite number of points in time. For simplicity, suppose that $M+1$ equidistant price observation are available each day, $p(t-1), p(t-1 + \frac{1}{M}), \dots, p(t)$. The j^{th} within-day return is then simply defined by

$$r_{t,j} = p\left(t-1 + \frac{j}{M}\right) - p\left(t-1 + \frac{j-1}{M}\right), \quad j = 1, 2, \dots, M, \quad (2)$$

for a total M returns per day.

As discussed at length in, e.g., Andersen et al. (2002), the *realized variance*

$$RV_t = \sum_{j=1}^M r_{t,j}^2. \quad (3)$$

provides a natural measure of the daily *ex-post* variation. In particular, it is well known that for increasingly finer sampling frequencies, or $M \rightarrow \infty$, RV_t consistently estimates the total variation comprised of the integrated variance plus the sum of the squared jumps

$$\lim_{M \rightarrow \infty} RV_t = \int_{t-1}^t \sigma^2(s)ds + \sum_{k=1}^{N_t} \kappa_{t,k}^2, \quad (4)$$

⁶This particular notation is adopted from Basawa and Brockwell (1982). A common modeling assumption for the Lévy process is the Compound Poisson process, or rare jump process, where the jump intensity is constant and the jump sizes are independent and identically distributed.

where N_t denotes the number of within-day jumps on day t , and $\kappa_{t,k}$ refer to the size of the k^{th} such jump.

In order to separately measure the two components that make up the total variation, Barndorff-Nielsen and Shephard (2004) and Barndorff-Nielsen et al. (2005) first proposed the so-called *bipower variation* measure

$$BV_t = \mu_1^{-2} \left(\frac{M}{M-1} \right) \sum_{j=1}^M |r_{t,j-1}| |r_{t,j}|, \quad (5)$$

where $\mu_1 = \sqrt{2/\pi} \approx 0.7979$. Under reasonable assumptions, it follows that

$$\lim_{M \rightarrow \infty} BV_t = \int_{t-1}^t \sigma^2(s) ds, \quad (6)$$

so that BV_t consistently estimates the integrated variance even in the presence of jumps in the underlying price process. Thus, as such the contribution to the total variation coming from jumps may be estimated by $RV_t - BV_t$. Additional insight may be gained by considering the *relative jump* measure of Huang and Tauchen (2005),

$$RJ_t = \frac{RV_t - BV_t}{RV_t}, \quad (7)$$

or the proportion of the total variation due to jump(s).⁷ In the limit as $M \rightarrow \infty$, $RJ_t > 0$ only on days for which there are at least one jump, although for finite M sampling variation can occasionally result in $RJ_t < 0$.

2.2 The BN-S Jump Statistic

In order to formally test for the presence of jumps, we need the joint asymptotic distribution of RV_t and BV_t . Under the null hypothesis of no jumps, and fairly mild regularity conditions about the price process in (1), Barndorff-Nielsen and Shephard (2006) show that the joint distribution

⁷An equivalent ratio statistic, $-RJ_t$, was proposed and studied independently by Barndorff-Nielsen and Shephard (2006).

of RV_t and BV_t conditional on the volatility path is mixed normal for $M \rightarrow \infty$,

$$M^{\frac{1}{2}} \left[\int_{t-1}^t \sigma^4(s) ds \right]^{-\frac{1}{2}} \begin{pmatrix} RV_t - \int_{t-1}^t \sigma^2(s) ds \\ BV_t - \int_{t-1}^t \sigma^2(s) ds \end{pmatrix} \xrightarrow{\mathcal{D}} N \left(0, \begin{bmatrix} \nu_{qq} & \nu_{qb} \\ \nu_{qb} & \nu_{bb} \end{bmatrix} \right), \quad (8)$$

where $\nu_{qq} = 2$, $\nu_{qb} = 2$, and $\nu_{bb} = (\pi/2)^2 + \pi - 3 \approx 2.6090$.

Determination of the scale of $RV_t - BV_t$ in units of conditional standard deviation still requires an estimate of the integrated quarticity, $\int_{t-1}^t \sigma^4(s) ds$. The *tripower quarticity* defined by

$$TP_t = \mu_{4/3}^{-3} M \left(\frac{M}{M-2} \right) \sum_{j=3}^M |r_{t,j-2}|^{\frac{4}{3}} |r_{t,j-1}|^{\frac{4}{3}} |r_{t,j}|^{\frac{4}{3}}, \quad (9)$$

where $\mu_{4/3} = 2^{2/3} \Gamma(7/6) / \Gamma(1/2) \approx 0.8309$, provides such an estimator. In particular, it is possible to show that

$$\lim_{M \rightarrow \infty} TP_t = \int_{t-1}^t \sigma^4(s) ds, \quad (10)$$

even in the presence of jumps.⁸

A variety of asymptotically equivalent test statistics may be formed by studentizing a measures of the discrepancy between RV_t and BV_t . Meanwhile, as noted by Barndorff-Nielsen and Shephard (2004), a studentized version of the aforementioned ratio statistics, RJ_t , may be expected to perform particularly well as it largely mitigates the effects of level shifts in variance associated with time varying stochastic volatility. This conjecture is corroborated by the extensive Monte Carlo evidence in Huang and Tauchen (2005), which indicate that the test statistic,

$$z_t = \frac{RJ_t}{\sqrt{(\nu_{bb} - \nu_{qq}) \frac{1}{M} \max \left(1, \frac{TP_t}{BV_t^2} \right)}}, \quad (11)$$

closely approximates a standard normal distribution under the null hypothesis of no jumps, $z_t \xrightarrow{\mathcal{D}} N(0, 1)$, and also exhibits favorable power properties in comparison to other transformations based on the bivariate asymptotic distribution for RV_t and BV_t in (8).⁹ For simplicity, we will

⁸Other jump-robust estimators for the integrated quarticity based on the summation of adjacent returns raised to powers less than two where the powers sum to four are also possible; see Barndorff-Nielsen and Shephard (2004). However, the Monte Carlo evidence in Huang and Tauchen (2005) indicates that the tripower quarticity performs quite well.

⁹The denominator of the statistic in (11) also incorporates a Jensen's inequality type adjustment for the (asymptotic) relationship between TP_t and BV_t^2 .

refer to the statistic in (11) as just the BN-S z -statistic. In what follows we use the statistic with a 0.001 significance level as a simple nonparametric jump detection tool for a large cross-section of individual stocks and an equiweighted index from the same stocks.¹⁰

3 Data

3.1 Date Source and Sampling

Our original sample consists of all trades on 40 large capitalization stocks over the January 1, 2001 to December 31, 2005, five-year sample period.¹¹ While the BN-S jump detection scheme is based on the notion of $M \rightarrow \infty$, or ever finely sampled high-frequency returns, a host of practical market microstructure complications prevents us from sampling too frequently while maintaining the fundamental semimartingale assumption underlying equation (1). Ways in which to best deal with these complications and the practical choice of M are the subject of intensive ongoing research efforts; see, e.g., Ait-Sahalia et al. (2005), Bandi and Russell (2005), Barndorff-Nielsen et al. (2006), and Hansen and Lunde (2006). In the analysis reported on below, we simply follow most of the literature in the use of a coarse sampling frequency as a way to strike a reasonable balance between the desire for as finely sampled observations as possible on the one hand and the desire to not be overwhelmed by the market microstructure noise on the other. The volatility signature plots advocated by Andersen et al. (2000), as further detailed in the Data Appendix, suggest that a choice of $M = 22$, or 17.5 minute sampling, strikes such a balance and largely mitigates the effect of the "noise" for all of the 40 stocks in the sample.¹²

In addition to high-frequency returns for each of the individual stocks, we also construct an equiweighted portfolio comprised of the same 40 stocks. We will refer to this index as EQW in the sequel. It is noteworthy that at the 17.5 minute sampling frequency, the correlation between the return on EQW and the return on the exchange-traded SPY fund, which tracks the S&P

¹⁰We also experimented with the use of other significance levels, resulting in the same basic findings and empirical puzzles. Some of these additional results are included in the tables, but we only discuss them briefly in the main text.

¹¹The criteria used in selecting the 40 stocks are discussed in more detail in the Data Appendix. The specific ticker symbols for each of the stocks are included in many of the figures.

¹²For simplicity we decided to maintain the identical sampling frequency across all stocks. However, $M = 22$ is clearly a conservative choice for many of the stocks, and we also calculated the same statistics based on finer sampled high-frequency returns. Some of these additional results are briefly discussed below.

500, equals 0.93.¹³ As such, the return on the EQW index may reasonably be thought of as being representative of the return on the aggregate market, and we will sometimes refer to as the market portfolio for short.

3.2 An Illustrative Look at Proctor and Gamble (PG)

Before presenting the results for all of the 40 stocks and the EQW index, it is instructive to look at the different variation measures and the BN-S test for a single stock and a few specific days in the sample. To this end, Figure 2 shows the price (adjusted for stock splits) and returns for PG. While the price appears to be steadily increasing over the 1241 trading days in the 2001-2005 sample, the corresponding 27,302 high-frequency 17.5 minute returns are all seemingly scattered around zero. At the same time, the return plot clearly indicates the presence of volatility clustering. This is further underscored by Figure 3, which plots the RV_t and BV_t variation measures, the relative jump contribution RJ_t , along with the BN-S z -statistic in equation (11). Comparing the BN-S test statistic in the lower plot to the horizontal reference line for the 99.9 percent significance level also included in the plot, indicates that PG jumped at least once during the active part of the trading day on 17 days in the sample.

To further illustrate the working of the jump detection scheme, Figure 4 shows the intraday prices and returns for PG on March 26 and 27, 2001. At a first glance it appears as if the price evolves rather smoothly on the 26th, while it shows a rather sharp increase just before noon on the 27th. Nonetheless, the corresponding BN-S z -statistics equal to 3.63 and -0.20 for each of the two days, respectively, suggest at least one highly significant jump on the 26th and no jumps on the 27th. Importantly, the jump test statistic depends on both the magnitude of the largest price change(s) over the day and the overall level of the volatility for the day. As such, relatively small price changes may be classified as jumps on otherwise calm days, while apparent discontinuities may be entirely compatible with a continuous sample path process in a statistical sense on very volatility days. Indeed, looking at the plot of the returns in the lower panel, the overall level of the volatility was clearly much higher on the 27th than on the 26th, with the BV_t estimate for the continuous sample path variation for each of the two days equal to 5.11 and 1.04, respectively. Thus, while some of the statistically significant jumps identified by the BN-S test can rather easily be spotted by *ex-post* visual inspection of the intraday prices

¹³At the five-minute interval the correlation is 0.88.

that is not necessarily always the case.

4 The Puzzle

We begin by showing in Figure 5 the number days over the 2001–2005 sample period on which the BN-S z -statistics indicate that the EQW index and each of the 40 stocks comprising the index jumped. A day is classified as jump day if the z -statistic exceeds the critical value of the Gaussian distribution at the 0.001 significance level. Not surprisingly, the index is seen to jump on far fewer days than its components; the count for EQW is seven days while there are twenty two jump days on average for the individual stocks. Indeed, this finding may be seen as a simple consequence of basic portfolio theory. Stock-specific idiosyncratic jumps are diversified away in forming the index, so that the index can only jump less frequently than its components.

As a further robustness check, and to guard against non-synchronous trading effects in the EQW portfolio returns, Table 1 shows the number of detected jumps in the EQW index, the SPY cash index, and the average over the 40 stocks for a wide range of different sampling frequencies (M ranging from 10 to 385) and three significance levels (0.001, 0.01, and 0.05). There is obviously a fairly close agreement between the number of jumps detected in the EQW index and the SPY down to the five-minute level, after which microstructure effects appear to cause a divergence. Our finding that the index jumps less than its components on average is also robust, as the ratio between the number of jumps in the index to the average number of jumps in the individual stocks systematically remains below one for most sampling frequencies and significance levels. Again, this holds true for the EQW index as well as the exchange traded SPY fund.

The real puzzle arises when one tries to relate the jumps in the index back to the jumps in its components. The top panel of Figure 6 shows the estimated β on EQW for each of the 40 stocks. The β 's are all rather close to unity, which is not surprising for large-cap stocks. The β 's suggest that the returns on each stock moves approximately one-for-one with the index return, apart from the idiosyncratic component. In particular, large movements, or jumps, in the index should show up as large movements or jumps in the stock prices. On the basis of the estimated β 's, one would expect a reasonably high correlation between a measure of the likelihood of a jump in the index and a jump in the individual stock prices. We can think of the BN-S z -statistic

(11) as such a measure. The middle panel shows the 40 regression slope coefficients while the bottom panel of Figure 6 shows the 40 estimated correlations between the stock z -statistics and the EQW z -statistic. Contrary to expectations, the correlations are exceedingly low: rarely above 0.05 and frequently on the order of 0.01 or 0.005, while the regression slope coefficients are small and insignificant as well.

This low correlation between the jump-test z -statistics at the level of the individual stocks and the z -statistics for the EQW portfolio is further illustrated in Figure 7, which shows a scatter plot of the average BN-S z -statistics for the 40 individual stocks against the z -statistic for EQW for each of the 1241 days in the sample. The relationship is obviously very weak. Indeed, for those days on which the z -statistic for EQW is statistically significant at the 0.001 level, the average z -statistic across the 40 stocks is just about 1.0, far below the cutoff value for statistical significance at any commonly used level.¹⁴

As final way to assess the relationship, or the apparent lack thereof, between the jumps in the aggregate index and the individual stocks, Figure 8 plots the z -statistics for each of the 40 individual stocks for the six days in the sample with the largest EQW z -statistics. Very few of the individual z -statistics are significant at the 0.001 level, and in fact for some of the market jump days *none* of the individual z -statistics exceed the 0.001 level. Meanwhile, the effects of diversification imply that any jumps in a large equiweighted index must necessarily be common jumps, or cojumps. Hence, the only explanation for these findings is that there are cojumps that are not being detected by the tests when applied to the individual stocks.

More formally, recall from equations (4) and (6) that

$$RV_{i,t} \xrightarrow{M \rightarrow \infty} \int_{t-1}^t \sigma_i^2(s) ds + \sum_{k=1}^{N_{t,i}} \kappa_{i,t,k}^2,$$

$$BV_{i,t} \xrightarrow{M \rightarrow \infty} \int_{t-1}^t \sigma_i^2(s) ds,$$

where $\kappa_{i,t,k}$ refers to the k^{th} jump in stock i on day t , and $N_{t,i}$ denotes the number of jumps in

¹⁴As pointed out by Roel Oomen, under the null hypothesis of no jumps in *any* of the 40 stocks the distribution of the average of the individual z -statistics will not be standard normal, but rather normally distributed with a mean of zero and a variance of 1/40, so that any value for the average in excess of 0.489 would be deemed significant at the 0.001 level.

stock i on day t . Hence,

$$RV_{i,t} - BV_{i,t} \xrightarrow{M \rightarrow \infty} \sum_{k=1}^{N_{t,i}} \kappa_{i,t,k}^2.$$

Now consider the j^{th} within-day return on an equiweighted portfolio of n stocks

$$r_{EQW,t,j} = \frac{1}{n} \sum_{i=1}^n r_{i,t,j}.$$

Extending the arguments for the individual stocks, the realized variation for the EQW portfolio must satisfy

$$\begin{aligned} RV_{EQW,t} &= \sum_{j=1}^M \left(\frac{1}{n} \sum_{i=1}^n r_{i,t,j} \right)^2 \xrightarrow{M \rightarrow \infty} \\ &\frac{1}{n^2} \sum_{i=1}^n \int_{t-1}^t \sigma_i^2(s) ds + \frac{1}{n^2} \sum_{i=1}^n \sum_{\ell=1, \ell \neq i}^n \int_{t-1}^t \sigma_i(s) \sigma_\ell(s) ds + \\ &\frac{1}{n^2} \sum_{i=1}^n \sum_{k=1}^{N_{i,t}} \kappa_{i,t,k}^2 + \frac{1}{n^2} \sum_{i=1}^n \sum_{\ell=1, \ell \neq i}^n \sum_{k=1}^{N_t^*} \kappa_{i,t,k} \kappa_{\ell,t,k}, \end{aligned}$$

where the last sum only includes the N_t^* jumps (a random number) that occur simultaneously across all n stocks. Similarly,

$$\begin{aligned} BV_{EQW,t} &= \sum_{j=2}^M \left| \frac{1}{n} \sum_{i=1}^n r_{i,t,j-1} \right| \cdot \left| \frac{1}{n} \sum_{i=1}^n r_{i,t,j} \right| \xrightarrow{M \rightarrow \infty} \\ &\frac{1}{n^2} \sum_{i=1}^n \int_{t-1}^t \sigma_i^2(s) ds + \frac{1}{n^2} \sum_{i=1}^n \sum_{\ell=1, \ell \neq i}^n \int_{t-1}^t \sigma_i(s) \sigma_\ell(s) ds. \end{aligned}$$

Thus, it follows readily that

$$\begin{aligned} RV_{EQW,t} - BV_{EQW,t} &\xrightarrow{M \rightarrow \infty} \frac{1}{n^2} \sum_{i=1}^n \sum_{k=1}^{N_{t,i}} \kappa_{i,t,k}^2 + \frac{1}{n^2} \sum_{i=1}^n \sum_{\ell=1, \ell \neq i}^n \sum_{k=1}^{N_t^*} \kappa_{i,t,k} \kappa_{\ell,t,k} \\ &= \frac{1}{n} \overline{\kappa^2} + \frac{n-1}{n} \sum_{k=1}^{N_t^*} \overline{\kappa_{i,t,k} \kappa_{\ell,t,k}}. \end{aligned}$$

For n large, the first term becomes negligible and $\frac{n-1}{n} \approx 1$ so that

$$RV_{EQW,t} - BV_{EQW,t} \approx \sum_{k=1}^{N_t^*} \overline{co_{t,k}}, \quad (12)$$

where $\overline{co_{t,k}}$ denotes the k^{th} average cojump. In other words, in a large well diversified portfolio, jumps can only be caused by jumps that occur simultaneously across assets, or jumps that pervade the market.

5 Cojumps and Risk

5.1 The cp -Statistic

As discussed in the previous section, jumps in a large well-diversified market index must necessarily be due to cojumps in the individual stocks that make up the index. In contrast to the BN-S and other related univariate jump detection procedures that examine the stocks or the index one-by-one, more powerful identification procedures for cojumps may therefore be obtained by examining the cross co-movements among the individual stocks.

To this end, consider the *cross product* statistic defined by the normalized sum of the individual high-frequency returns for each within-day period,

$$cp_{t,j} = \frac{1}{2n(n-1)} \sum_{i=1}^{n-1} \sum_{\ell=i+1}^n r_{i,t,j} r_{\ell,t,j}, \quad j = 1, 2, \dots, M. \quad (13)$$

The cp -statistic provides a direct measure of how closely the stocks move together. It is entirely analogous to a U -statistic. For each day in the sample we have a realization $\{cp_{t,j}\}_{j=1}^M$ of length M . Summing the $cp_{t,j}$ statistics across day t , it follows readily that

$$cp_t = \sum_{j=1}^M cp_{t,j} = \frac{1}{4(n-1)} \left[nRV_{EQW,t} - \frac{1}{n} \sum_{i=1}^n RV_{i,t} \right]. \quad (14)$$

Hence, from reasoning much like that underlying the BN-S bipower variation statistic in (5), we can expect the cp -statistic to be reasonably insensitive to idiosyncratic jumps in the individual stocks. At the same time, the statistic is clearly very sensitive to cojumps and it will assume a

large positive value whenever the ensemble of returns takes a large (positive or negative) move together.

Meanwhile, the cross correlation among the diffusive components of the returns implies that even in the absence of jumps, the cp -statistic does not have mean zero. Furthermore the overall scale of the statistic fluctuates randomly with the general level of stochastic volatility in the market. We thus studentize the statistic as

$$z_{cp,t,j} = \frac{cp_{t,j} - \overline{cp}_t}{s_{cp,t}} \quad j = 1, 2, \dots, M, \quad (15)$$

where

$$\overline{cp}_t = \frac{1}{M} cp_t = \frac{1}{M} \sum_{j=1}^M cp_{t,j}$$

and

$$s_{cp,t} = \sqrt{\frac{1}{M-1} \sum_{j=1}^M (cp_{t,j} - \overline{cp}_t)^2}.$$

The presumption here is that the location and scale are approximately constant within-days, but varies across days. This, of course, is entirely consistent with the idea of slowly varying stochastic return volatility. Also, as indicated by the average of the within-day serial correlation coefficients depicted in Figure 9, the $cp_{t,j}$ -statistics are essentially serially uncorrelated for $j = 1, 2, \dots, M$. Hence, we simply standardize each of the within-day cp -statistics by their corresponding daily sample standard deviation, $s_{cp,t}$.

In order to actually use the normalized z_{cp} -statistic in (15) as a measure or test for cojumps, we need its distribution under the null of no jumps. Since the cp -statistic in (13) is itself the average of $n(n-1)/2$ random variables, where $n = 40$ in our data set, one might expect the distribution of the z_{cp} -statistic to be approximately standard $N(0, 1)$. In particular, Borovkova et al. (2001) give a central limit theorem for U -statistics for dependent data that would apply to the z_{cp} -statistic for $n \rightarrow \infty$ if the dependence (mixing conditions) among the individual stocks was sufficiently weak. However, the cross-correlation among stock returns is unlikely to satisfy the necessary conditions. For instance, it is fairly easy to show that under a simple one-factor representation for the returns, the cp -statistic will, for large n , equal $\mu_\beta^2 \chi^2(1)$, where μ_β^2 is the square of the average (arithmetic) β . The average β is arguably unity, which gives an asymptotic $\chi^2(1)$ representation for the cp -statistic. However, the simple $\chi^2(1)$ representation does not

hold in empirically more realistic multi-factor situations, and as such is of limited practical use. Lacking an appropriate approximating asymptotic distribution for the z_{cp} -statistic, we thus adopt a straightforward bootstrap procedure to get its distribution under the null of no jumps. This approach is a tractable compromise to employing the powerful theory of Barndorff-Nielsen and Shephard (2003), which appears difficult to implement empirically and is deferred to future work.

In particular, simulated values of the cp -statistic are readily computed along realizations of a 40×1 diffusion with zero drift and covariance matrix determined by the unconditional covariance matrix of the within-day returns. More specifically, we simulate realizations of length equal to the sample size (22 steps per day for 1241 days) replicated 1,000 times. This scheme thus generates just over 27.3 million simulated values under the null of no jumps. The top panel of Figure 10 shows this bootstrapped probability density of the z_{cp} -statistic. The distribution is evidently highly non-Gaussian with a strong right skew. The 99.9 percent quantile, or critical value for a 0.001 level test for no cojumps, equals 4.145.

As a check on the sensitivity of this critical value to the covariance structure of the returns, the number of stocks, and the number of within-day returns, we recomputed the bootstrap distribution using equicorrelated returns based on different values of $\bar{\rho}$, n and M . The results displayed in Table 2 reveal that the critical value is quite insensitive to the level of correlation among the returns and to the number of stocks. It is, however, somewhat sensitive to the number of within-day returns, M . Intuitively, using the daily sample mean and standard deviation in studentizing the z_{cp} -statistic in equation (15) in place of the population quantities, causes the distribution of the statistic to become more concentrated around zero.¹⁵ All together, it appears as if the bootstrap distribution shown in the top panel of Figure 10 accurately captures the distribution of the z_{cp} -statistic in the absence of jumps for the actual high-frequency panel analyzed here.

¹⁵If the returns were perfectly correlated, the cp -statistic would be distributed as $[\chi^2(1) - 1]/\sqrt{2}$ if population quantities were used to studentize; the 0.001 critical value would be 6.949. On the other hand, if $\rho = 0.995$, $M = 22$, and sample values are used to studentize, the bootstrap critical value equals just 4.266. Of course, for $\rho = 0.995$ and larger values of M the bootstrap critical value is much closer to the limiting value of 6.949.

5.2 Resolution of the Puzzle

Contrasting the bootstrap distribution in the top panel in Figure 10 to the empirical distribution of the z_{cp} -statistic in (15) computed from our panel of 40 stocks reveals quite striking evidence for cojumps. In particular, the corresponding Q - Q plot shown in the bottom panel in Figure 10 clearly deviates from the 45-degree line that should obtain in the absence of jumps. The empirical distribution is evidently right-shifted relative to the null distribution, and many of the sample z_{cp} -statistics would be judged as statistically significant at most any commonly use significance level.

This strong empirical evidence for a significant number of cojumps is further corroborated by Figure 11, which displays a scatter plot of the return on the EQW index against the z_{cp} -statistics. There are a total of $1241 \times 22 = 27,302$ points in the figure. As noted above, the 99.9 percent bootstrap quantile of the distribution of the z_{cp} -statistic equals 4.145, and so we should expect only about 32 points in Figure 11 to lie to the right of this cutoff. There are, however, far many more points than that to the right of the cutoff, although in many instances the associated return on the EQW index is only moderately large in magnitude. In other words, it appears that many modest-sized cojumps simply go undetected by the BN-S test statistic when applied to returns individually.

In order to reconcile our finding of many more cojumps based on the z_{cp} -statistic compared to that suggested by the BN-S statistic applied stock-by-stock, it is instructive to consider the magnitude of the background noise and the average size of the cojumps. Since the values of the β 's are all about unity, the cojumps are certainly present in the individual stocks. However, they are masked by the high volatility associated with the idiosyncratic noise component and therefore remain largely undetected by the BN-S z -statistic. Similarly, because so many of the cojumps are only of modest size, they also go undetected by the BN-S z -statistic applied to the EQW return using a stringent 0.001 significance level.

We believe the above conclusions concerning the presence of many modest sized non-diversifiable cojumps to be quite persuasive. Nonetheless, to further buttress the empirical findings, we present the results from some additional Monte Carlo simulations designed to assess the performance of the cp -test vis-a-vis the BN-S procedure in situations with cojumps. Before we discuss these additional simulation results, however, we first discuss how the high-frequency data also reveals the existence of a strong intraday pattern in the occurrence of cojumps.

5.3 Intraday Patterns

There is a long history dating back at least to Wood et al. (1985) and Harris (1986) documenting the existence of a distinct U-shaped pattern in equity return volatility over the trading day; i.e., volatility tend to be high at the open and close of trading and low in the middle of the day. This same general pattern, albeit more tilted towards the open, is also apparent from Figure 12 and the plot of the unconditional variance and bipower variation for the EQW returns over the trading day.

More specifically, we compute the unconditional realized variance for tick j as

$$\frac{1}{T} \sum_{t=1}^T r_{EQW,t,j}^2, \quad j = 1, \dots, M,$$

while the unconditional bipower variation is computed as

$$\frac{1}{T} \sum_{t=1}^T |r_{EQW,t,j-1}|^{\frac{1}{2}} |r_{EQW,t,j}| |r_{EQW,t,j+1}|^{\frac{1}{2}}, \quad j = 2, \dots, M-1.$$

Figure 12 shows these unconditional variance measures at three increasingly finer sampling frequencies: $M = 22$, 77 , and 385 , corresponding to 17.5, five, and one minute, respectively.¹⁶ The unconditional realized variance systematically lies above the unconditional bipower variation over the entire day, thus reflecting the existence of jumps across the day. Interestingly, however, the plot also reveals a sharp rise in the realized variation relative to the bipower variation at 10am EST, corresponding to the time of the release of several regularly scheduled macroeconomic news announcements.¹⁷ This therefore indirectly suggests that some of the cojums may be associated with these types of systematic news. This is also consistent with the work of Andersen et al. (2003) and Andersen et al. (2007) among others, which document a significant response in high-frequency financial market prices to surprises in macroeconomic announcement immediately after the release of the news. Moreover, Andersen and Bollerslev (1998) among others have previously noted a sharp increase in the average total intraday volatility at the exact time of important macroeconomic news announcements. What is particularly noteworthy

¹⁶For visual comparisons, we extended the bipower variation directly to the left for $j = 1$ and to the right for $j = M$.

¹⁷There is also a smaller less pronounced peak at 1:15pm. However, other computations by Peter Van Tassel, not shown here, suggest that this early afternoon peak is driven by a few extreme observations. Also, the pattern in Figure 12 for the EQW portfolio essentially mirrors that for the SPY.

here, however, is the much less dramatic increase in the average within-day bipower variation measure, in turn attributing most of the variation at that specific time-of-day to jumps.

Instead of averaging all of the individual returns before calculating the intraday variation measures as in Figure 12, Figure 13 shows a similar plot in which the two variation measures are first computed on a stock-by-stock basis and then averaged across all of the 40 stocks in the panel. This explicitly excludes the effect of diversification, and as results the vertical scale of Figure 13 is much larger than that of Figure 12. More importantly, however, comparing the general shape between the two sets of pictures, the increase in the within-day variation at 10am is much less apparent for the individual stock averages. The relative importance of jumps also appears to be much more evenly distributed across the entire day, and as such lend further credence to our hypothesis of the cojumps drowning in the firm-specific variation inherent in the individual stock returns.

Meanwhile, the z_{cp} -statistic in (15) that we used in testing for cojumps was based on the assumption of constant scale within the day. The average within-day patterns in Figure 13 clearly seem to violate this assumption. At the same time, there is no obvious right way to adjust the individual returns for the intraday volatility patterns, as the relative importance of the jump variation to the total variation may be changing over the day. Fortunately, that does not seem to matter much for our main conclusions. In particular, the cojump statistics depicted in the previous Figure 11 were based on the raw unadjusted returns ignoring the intraday pattern. As a robustness check we redid the same calculations in which we scaled the return for the i^{th} stock over the j^{th} interval by the reciprocal of the square root of the unconditional bipower variation of the stock for that time-interval. This adjustment is extreme, in that it deflates returns near the beginning and end of the day while inflating returns in the middle of the day under the implicit presumption that the share of the jump variance remains constant over the day. Nonetheless, the resulting Figure 14 depicting the returns on the EQW index against the adjusted z_{cp} -statistics, is so similar to the original Figure 11 that we conjecture any reasonable adjustment for the intraday pattern would result in the same basic conclusions.

All-in-all, we believe the conclusions in the preceding two subsections to be quite persuasive. However, to further buttress the findings, we next present the results from some additional Monte Carlo simulations specifically designed to illustrate the ability of the cp -test to actually detect true cojumps.

5.4 Further Validation of the cp -Statistic

Our main conclusions hinge on the argument that the z_{cp} -statistic is more sensitive to cojumps than the BN-S z -statistic applied stock-by-stock because it explicitly utilizes the cross-sectional information. To investigate this hypothesis further we use another bootstrap-type procedure where we take the observed data set as given, sprinkle in additional simulated jumps, and then recompute the jump test statistics.

Specifically, for idiosyncratic jumps we simulated 40 independent Gaussian Compound Poisson processes with intensity λ_i and jump sizes $N(0, \sigma_{J,i}^2)$, while for the common jumps we simulated one Gaussian Compound Poisson process with intensity λ and jump sizes $N(0, \sigma_J^2)$. The idiosyncratic jumps are then added to the actually observed within-day 17.5 minute returns for each of the individual stocks as are the common jumps multiplied by the stock's estimated β . This in turn yields 40 new returns series with the additional simulated jumps scattered throughout the sample. From these new series, we then recompute the EQW returns, the BN-S z -statistics, and the z_{cp} -statistics. This whole process is replicated 1,000 times and the outcomes averaged across the replications. Since the baseline observed data already contains jumps, the simulations reveal the *incremental* strength, or power, of the tests to the different types of added jumps.

The findings for different jump sizes holding the intensities constant are readily described and intuitively plausible. In the absence of any common jumps, as the idiosyncratic jump sizes increase from zero percent to 1.50 percent, the jump detection rate of the BN-S z -statistic applied to the individual stocks gradually increases. However, it remains unchanged when applied to the EQW return, as does the detection rate of the z_{cp} -statistic. The idiosyncratic jumps are effectively diversified away. On the other hand, if the idiosyncratic jumps are left out while the size of the common jump increases from zero percent to 1.50 percent, the detection rates of both the BN-S z -statistics and z_{cp} -statistics increase, with the z_{cp} -statistic being slightly more sensitive to larger jump sizes. Still, the contrasts between the two tests appear rather small, as they both fairly easily detect the occasional rare common jump.

The most revealing contrasts between the statistics is obtained by considering their ability to detect a relatively small common jump. Specifically, Figure 15 shows the jump detection rates for a simulation experiment in which the idiosyncratic jumps are left out, the common jump size is set to 0.25 percent, and the intensity λ of the common jumps increases from zero to 200 per

year. For the z_{cp} -statistic, there are $M = 22$ within-day values, so any day on which the statistic is statistically significant is classified as a jump day; in the observed data set, there are no days on which more than one statistically significant z_{cp} -statistic occurs. For the BN-S z -statistic, the figure shows the number of detected jumps using the BN-S z -statistic applied to the EQW returns as well as the average number of jumps detected from applying it one-by-one across the panel of 40 stocks. Note that for $\lambda = 0$ there are more jumps detected than the nominal size of the tests as the baseline data contains jumps. Nonetheless, as clearly seen from the figure, the detection rate of the z_{cp} -statistic increases more sharply with λ than do the detection rates based on the BN-S statistic applied stock-by-stock. This directly confirms the intuition that utilizing the cross covariance structure provides more information on the non-diversifiable cojumps than does applying the standard univariate tests to either the EQW index or to the individual stocks one-by-one.

6 Conclusion

Using popular high-frequency data based jump detection procedures we document an apparent disconnect in the number of significant jumps in individual stock and aggregate index returns, with jumps appearing more than three times as likely at the individual stock level. The fact that the index jumped less, on average, than the individual stocks is a simple reflection of diversification of idiosyncratic jumps. However, we also find that the values of the jump test statistics for the individual stocks are largely uncorrelated with the values of the test statistic for an index constructed from the very same stocks. This lack of correlation is a genuine anomaly in view of the fact that all of the stocks have a β of about unity with respect to the index, and that an index can jump only if one or more of its components jumps.

The resolution of the puzzle lies in the presence of many moderate-sized cojumps shared among the stocks. These cojumps remain undetected by the standard statistics at the level of the individual stocks because of the large background noise. Since many of the cojumps are only of moderate size, they also remain largely undetected at the index level. To more effectively detect cojumps we develop a new *cross product* statistic, termed the cp -statistic, that directly uses the cross-covariation structure of the high-frequency returns. Employing this statistic we successfully detect many modest-sized cojumps, in turn confirming our resolution of the puzzle.

Documenting the presence of cojumps and understanding their economic determinants and dynamics are crucial from a risk measurement and management perspective. Basic portfolio theory implies that the only kind of jumps that can carry a risk premium are a non-diversifiable cojumps. Measuring the risk premium on cojumps is far beyond the scope of the present paper. However, using index-level data Todorov (2006) makes progress towards separating the aggregate jump risk premium from the continuous volatility risk premium and understanding its dynamics. The ideas and techniques developed here may prove especially useful in future work along these lines.

Data Appendix

We initially selected the 50 most actively traded stocks on the New York Stock Exchange (NYSE) according to their ten-day trading volume (number of shares) at the beginning of June 2006. Of these 50 stocks, we were able to successfully download reliable high-frequency prices for 40. The ticker symbols for these 40 stocks are included in many of the figures.

Data Source and Cleaning

Data on all completed trades were obtained from the Trade and Quote Database (TAQ) available via the Wharton Research Data Services (WRDS). This includes trades from all North American exchanges as well as over-the-counter trades. Each exchange has its own distinct market structure which might affect the structure of observed prices. Hence, to homogenize the data, we decided to only consider trades on the NYSE. The NYSE also accounts for the majority of the trades for all of the stocks in the sample.

Our sample covers the period from January 1st 2001 to December 31st 2005. Trading frequency increased significantly in the late 90s, and by the end of 2001 all NYSE listed stocks had moved from fractional to decimalized trading, in turn allowing for the extraction of highly reliable high-frequency prices. Illogical data values such as time stamp errors (e.g., hour #25, minute #78, month #43 and year #3001) and negative prices are removed from the data. All-in-all, these errors represent a relatively small number of data points. We also exclude trades that occur outside of 9.30am and 4pm, as well as days with only partial trading. Examples of such days are September 11, 2001, and certain holidays when the NYSE is only open for part of the day. A listing of all of these dates is available on the NYSE web-site.

Because of the unusual activity associated with trading at the beginning of each day, we start our intraday sampling at 9.35am, five-minutes after the market officially opens. This ensures a more homogenous trading and information gathering mechanism for all of the prices. The price series are sampled every 30-seconds using a slightly modified version of the previous tick method from Dacorogna et al. (2001). The previous tick method simply fixes the time where prices are ideally sampled at regular intervals and selects a completed trade prior to the time should there be no trade at that particular time. For instance, a trade completed at 9:34:58 is used in place of 9:35:00 when there is no actual trade at 9:35:00. In this case, there is therefore a 2-seconds backtrack, as defined by the time difference between the ideal sampling time and the actual sampling time. The first trade of the day is used if there are no prior trades on that day.

With 30-seconds sampling from 9.35am to 4.00pm this leaves us with 771 prices per day. Also, the sample period from January 1st 2001 through December 31st 2005 consists of 1241 normal trading days, for a total of close to a million transaction prices for each of the 40 individual stocks.

The raw high-frequency prices invariably contains a number of mis-recordings and other data errors. In some cases these errors are obvious by visual inspection of time series plots of the data, sometimes they are not. Thus, in addition to manually inspecting and correcting the data, we also employed a threshold filter of 1.5 percent, which appears to work well for removing and cleaning the remaining data errors at the 30-seconds sampling interval.

Sampling Frequency

The statistics used in the paper formally becomes more accurate as the sampling frequency increases. However, as noted in the main text of the paper, there is a limit to how finely we can sample the price process while maintaining the basic underlying semimartingale assumption as a host of market microstructure influences start to materially affect the observed price changes; most importantly features having to do with specifics of the trading mechanism, Black (1976) and Amihud and Mendelson (1987), and discreteness of the data, Harris (1990, 1991). As discussed at length in Hansen and Lunde (2006), the design of new procedures and "optimal" ways in which to deal with these complications is currently the focus of extensive research efforts. Rather than employing any of these more advanced procedures, in the analysis reported on here, we simply rely on the volatility signature plots proposed by Andersen et al. (2000) as an easy-to-implement procedure for choosing the highest possible sampling frequency so that the realized variation measures remain unbiased for the unconditional daily variance.

The corresponding plots for each of the 40 stocks in Figure 16 suggest that by sampling $M = 22$ times per day, or equivalently by using 17.5 minute returns in the construction of the realized variation measures, the market microstructure influences have essentially ceased and the plots become flat. Of course, the plots also suggest that for many of the stocks we could safely sample more frequently. However, for simplicity we decided to maintain the same sampling frequency across all 40 stocks throughout the entire sample. Of course, the use of identically sampled high-frequency returns across all stocks also facilitate the construction of the equiweighted index returns. Importantly, the use of $M = 22$ also involves relatively little interpolation in the construction of the equidistant 17.5 minute returns. The median backtrack

from the previous tick method for each of the 40 stocks depicted in Figure 17 is just about 6.5 seconds, and for none of the stocks does the median backtrack exceed 12 seconds.

References

- Ait-Sahalia, Y. and J. Jacod (2006). Testing for Jumps in a Discretely Observed Process. Working paper, Princeton University.
- Ait-Sahalia, Y., P. Mykland, and L. Zhang (2005). How Often to Sample a Continuous Time Process in the Presence of Market Microstructure Noise. *Review of Financial Studies* 18, 351–416.
- Amihud, Y. and H. Mendelson (1987). Trading Mechanisms and Stock Returns: An Empirical Investigation. *Journal of Finance* 42, 533–553.
- Andersen, T., B. Bollerslev, F. Diebold, and P. Labys (2000). Great Realizations. *Risk*.
- Andersen, T., B. Bollerslev, F. Diebold, and C. Vega (2003). Micro Effects of Macro Announcements: Real-Time Price Discovery in Foreign Exchange. *American Economic Review* 93, 38–62.
- Andersen, T., B. Bollerslev, F. Diebold, and C. Vega (2007). Real-Time Price Discovery in Stock, Bond and Foreign Exchange Markets. *Journal of International Economics*, forthcoming.
- Andersen, T. and T. Bollerslev (1998). Deutschemark-Dollar Volatility: Intraday Activity Patterns, Macroeconomic Announcements, and Longer Run Dependencies. *Journal of Finance* 53, 219–265.
- Andersen, T., T. Bollerslev, and F. Diebold (2002). *Parametric and Nonparametric Volatility Measurement*. In Handbook of Financial Econometrics. Amsterdam, North Holland: (eds.) Yacine Ait-Sahalia and Lars Peter Hansen.
- Andersen, T., T. Bollerslev, and F. Diebold (2007). Roughing it Up: Including Jump Components in the Measurement, Modeling and Forecasting of Return Volatility. *Review of Economics and Statistics* 89, forthcoming.
- Bandi, F. and J. Russell (2005). Microstructure Noise, Realised Volatility, and Optimal Sampling. Working paper, Graduate School of Business, University of Chicago.

- Barndorff-Nielsen, O., S. Graversen, J. Jacod, M. Podolskij, and N. Shephard (2005). A Central Limit Theorem for Realised Power and Bi-power Variations of Continuous Semimartingales. Working paper, Nuffield College, Oxford University.
- Barndorff-Nielsen, O., P. Hansen, A. Lunde, and N. Shephard (2006). Designing Realised Kernels to Measure the ex-post Variation of Equity Prizes in the Presence of Noise. Working paper, Nuffield College, Oxford University.
- Barndorff-Nielsen, O. and N. Shephard (2003). Measuring the Impact of Jumps in Multivariate Price Processes Using Bipower Covariation. Working paper, Nuffield College, Oxford University.
- Barndorff-Nielsen, O. and N. Shephard (2004). Power and Bipower Variation with Stochastic Volatility and Jumps. *Journal of Financial Econometrics* 2, 1–37.
- Barndorff-Nielsen, O. and N. Shephard (2006). Econometrics of Testing for Jumps in Financial Economics using Bipower Variation. *Journal of Financial Econometrics* 4, 1–30.
- Basawa, I. and P. Brockwell (1982). Non-Parametric Estimation for Non-Decreasing Lévy Processes. *Journal of the Royal Statistical Society Series B* 44, 262–269.
- Black, F. (1976). Noise. *Journal of Finance* 41, 529–543.
- Borovkova, S., R. Burton, and H. Dehling (2001). Limit Theorems for Functionals of Mixing Processes with Applications to U-Statistics and Dimension Estimation. *Transactions of the American Mathematical Society* 353, 4261–4318.
- Dacorogna, M., R. Gencay, U. Muller, R. Olsen, and O. Pictet (2001). *An Introduction to High Frequency Finance*. San Diego: Academic Press.
- Eraker, B., M. Johannes, and N. Polson (2003). The Impact of Jumps in Volatility and Returns. *Journal of Finance* 58, 1269–1300.
- Fan, J. and Y. Wang (2006). Multi-scale Jump and Volatility Analysis for High-Frequency Financial Data. Working paper, Princeton University.

- Hansen, P. and A. Lunde (2006). Realized Variance and Market Microstructure Noise. *Journal of Business and Economic Statistics* 24, 127–161.
- Harris, L. (1986). A Transaction Data Study of Weekly and Intradaily Patterns in Stock Returns. *Journal of Financial Economics* 16, 99–117.
- Harris, L. (1990). Estimation of Stock Variance and Serial Covariance from Discrete Observations. *Journal of Financial and Quantitative Analysis* 25, 291–306.
- Harris, L. (1991). Stock Price Clustering and Discreteness. *Review of Financial Studies* 4, 389–415.
- Huang, X. and G. Tauchen (2005). The Relative Contributions of Jumps to Total Variance. *Journal of Financial Econometrics* 3, 456–499.
- Jacod, J. and V. Todorov (2007). Testing for Common Arrivals of Jumps for Discretely Observed Multidimensional Processes. Working paper, Duke University.
- Jiang, G. and R. Oomen (2006). A New Test for Jumps in Asset Prices. Working paper, University of Warwick.
- Lee, S. and P. Mykland (2007). Jumps in Financial Markets: A New Nonparametric Test and Jump Dynamics. *Review of Financial Studies*, forthcoming.
- Mancini, C. (2004). Estimation of the Characteristics of the Jumps of a General Poisson-Diffusion Model. *Scandinavian Actuarial Journal*, 42–52.
- Todorov, V. (2006). Variance Risk Premium Dynamics. Working paper, Duke University.
- Wood, R., T. McInish, and J. Ord (1985). An Investigation of Transaction Data for NYSE Stocks. *Journal of Finance* 25, 723–739.

Table 1: Jump Counts At Different Sampling Frequencies

	Frequency		Jump Count			Ratios	
	Minutes	M	EQW	SPY	STK	EQW/STK	SPY/STK
0.001 Level	38.5	10	2	4	9.0	0.22	0.44
	35.0	11	5	7	11.7	0.43	0.60
	27.5	14	9	8	15.8	0.57	0.51
	17.5	22	7	11	22.1	0.32	0.50
	11.0	35	16	12	26.1	0.61	0.46
	7.0	55	33	24	31.5	1.05	0.76
	5.5	70	27	19	34.7	0.78	0.55
	5.0	77	32	38	36.7	0.87	1.04
	3.5	110	46	80	53.9	0.85	1.48
	2.5	154	46	100	87.2	0.53	1.15
	1.0	385	100	445	513.1	0.19	0.87
	38.5	10	48	35	63.0	0.76	0.56
	35.0	11	47	47	66.2	0.71	0.71
	27.5	14	53	43	68.7	0.77	0.63
	17.5	22	38	45	71.8	0.53	0.63
0.01 Level	11.0	35	58	55	88.7	0.65	0.62
	7.0	55	77	78	90.6	0.85	0.86
	5.5	70	83	88	96.7	0.86	0.91
	5.0	77	112	123	101.1	1.11	1.22
	3.5	110	122	170	133.7	0.91	1.27
	2.5	154	136	219	192.0	0.71	1.14
	1.0	385	205	591	717.1	0.29	0.82
	38.5	10	137	132	177.5	0.77	0.74
	35.0	11	156	158	180.6	0.86	0.87
	27.5	14	129	131	174.8	0.74	0.75
	17.5	22	110	121	175.5	0.63	0.69
	11.0	35	172	154	186.1	0.92	0.83
	7.0	55	202	208	204.0	0.99	1.02
	5.5	70	208	215	214.3	0.97	1.00
	5.0	77	255	250	220.2	1.16	1.14
0.05 Level	3.5	110	267	329	272.6	0.98	1.21
	2.5	154	290	398	355.8	0.82	1.12
	1.0	385	342	755	903.0	0.38	0.84

Table 2: 99.9 Percent Quantile of the Bootstrapped z_{cp} Distribution

	Average correlation $\bar{\rho}$	Quantile $Q_{0.999}$
Data Based		
$n = 40, M = 22$	0.35	4.15
Equicorrelated		
$n = 40, M = 22$	0.00	4.12
	0.10	4.14
	0.50	4.15
	0.80	4.15
$n = 20, M = 22$	0.00	4.09
	0.10	4.14
	0.50	4.15
	0.80	4.15
$n = 40, M = 78$	0.00	5.83
	0.10	5.89
	0.50	5.89
	0.80	5.89
$n = 20, M = 78$	0.00	5.76
	0.10	5.87
	0.50	5.89
	0.80	5.89

Figures

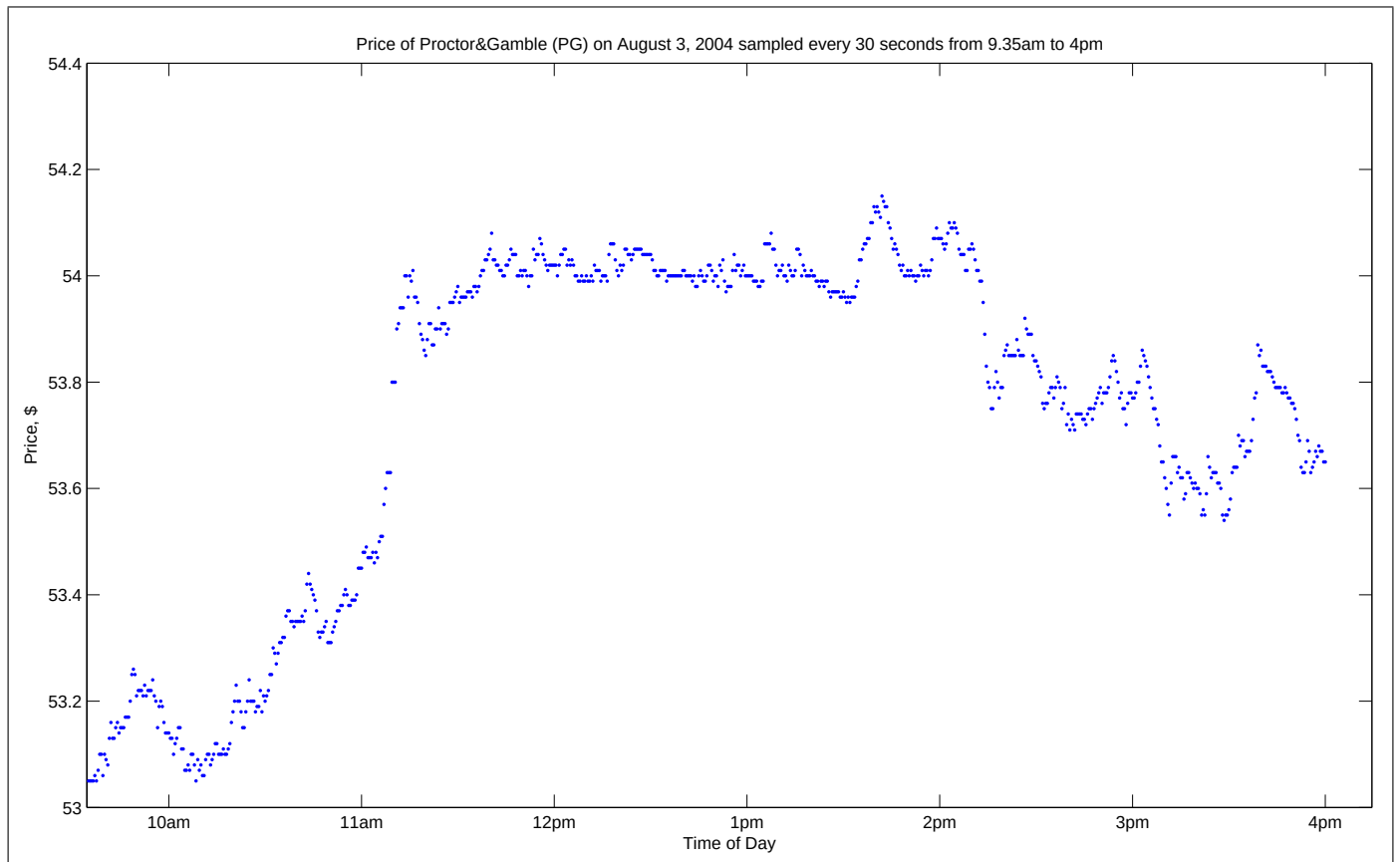


Figure 1: Price of PG sampled once every 30 seconds on August 3, 2004.

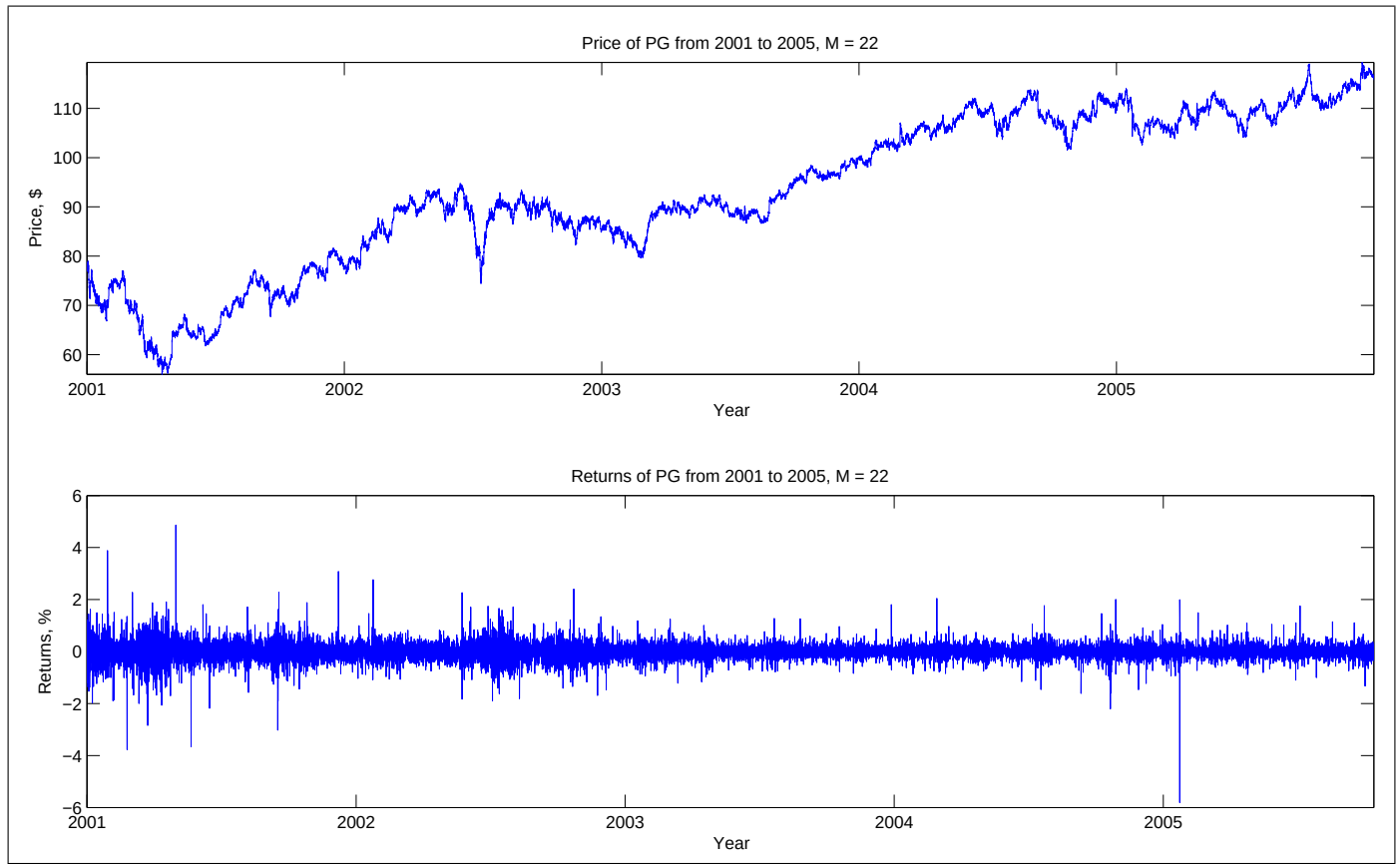


Figure 2: Price and Returns of PG from 2001 to 2005, $M = 22$.

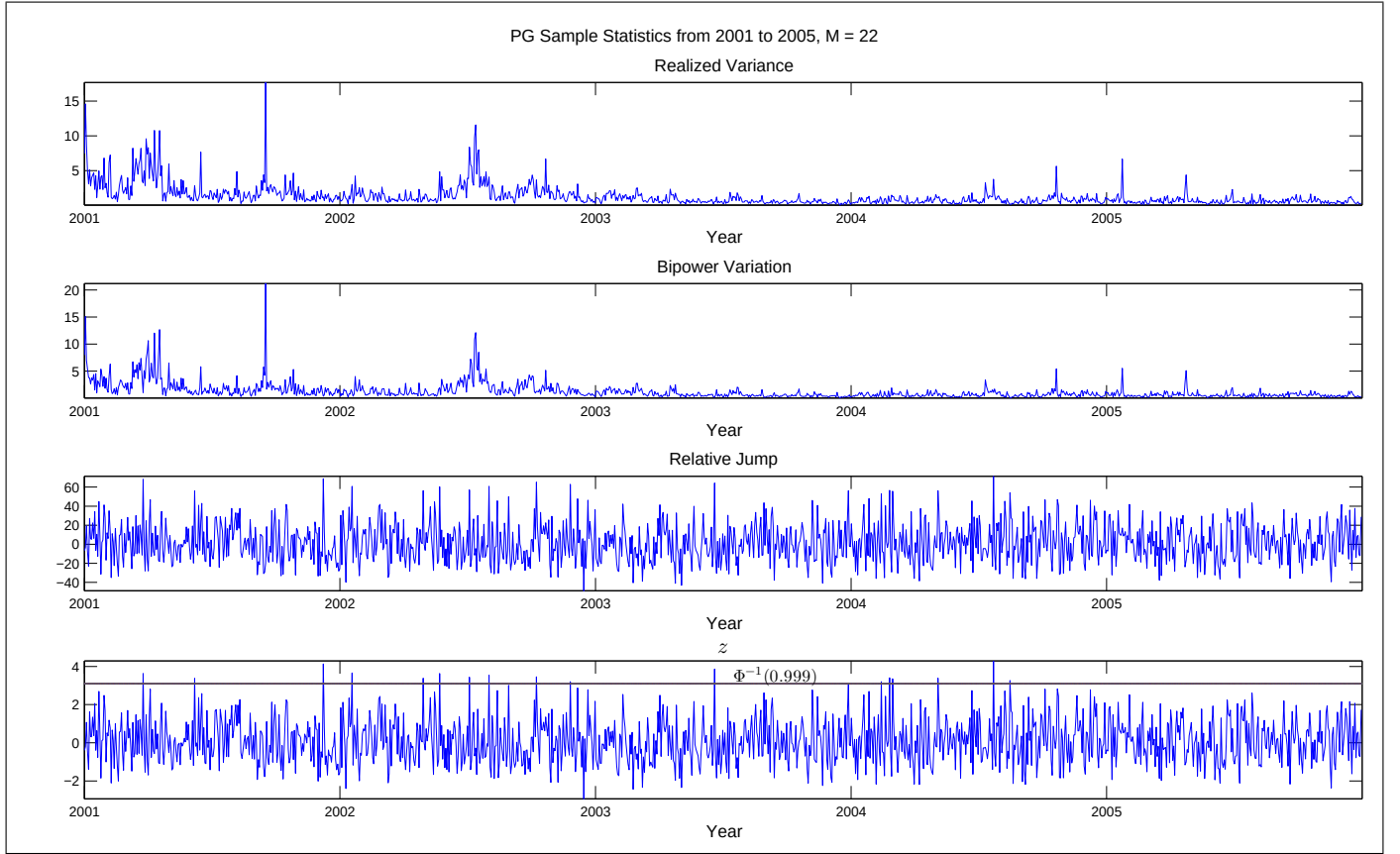


Figure 3: (Returns, Realized Variance, Bi-Power Variation, Relative Jump and z_t of PG from 2001 to 2005, M = 22.

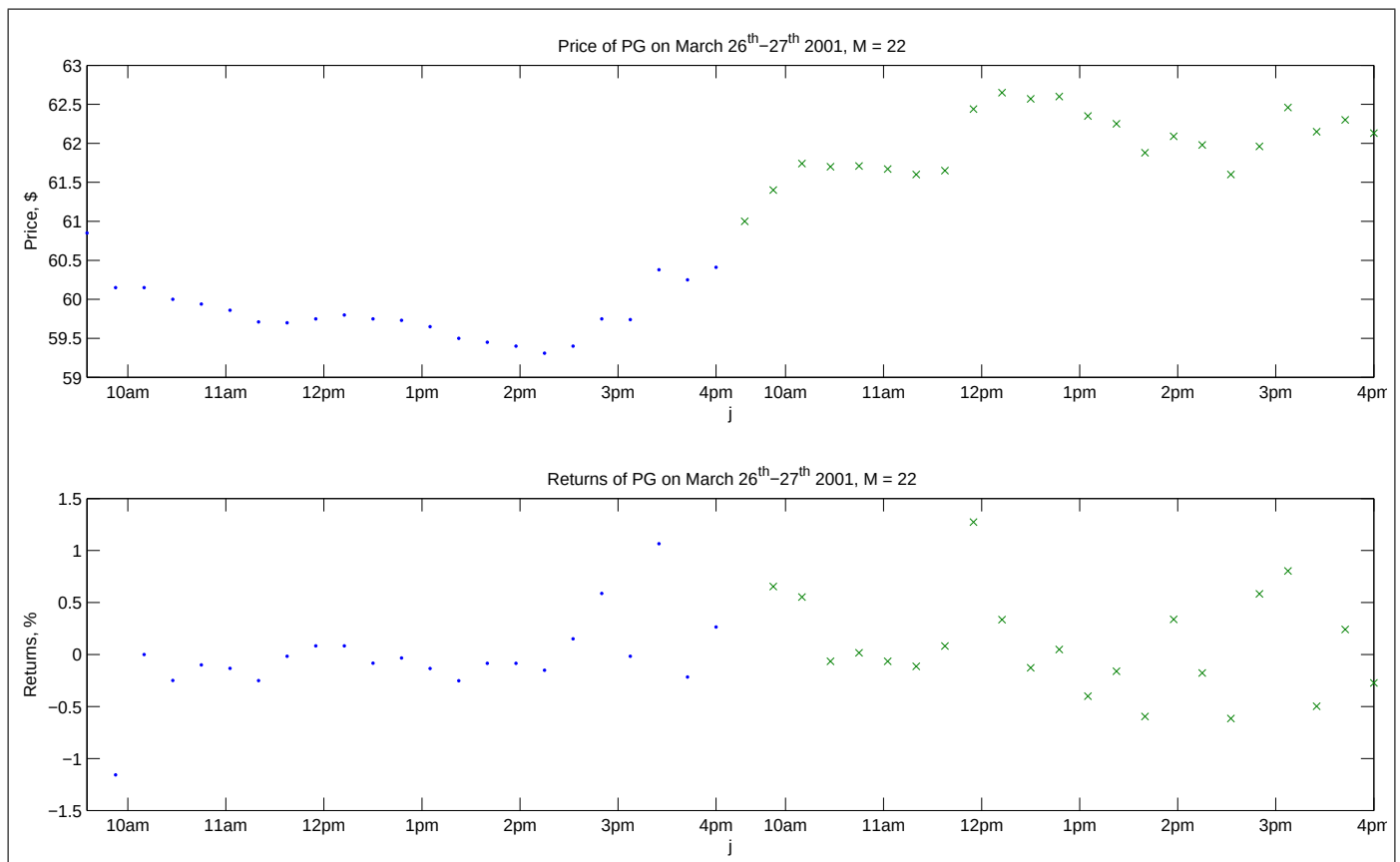


Figure 4: Price and corresponding returns of Procter&Gamble on March 26 - 27, 2001, $M = 22$.

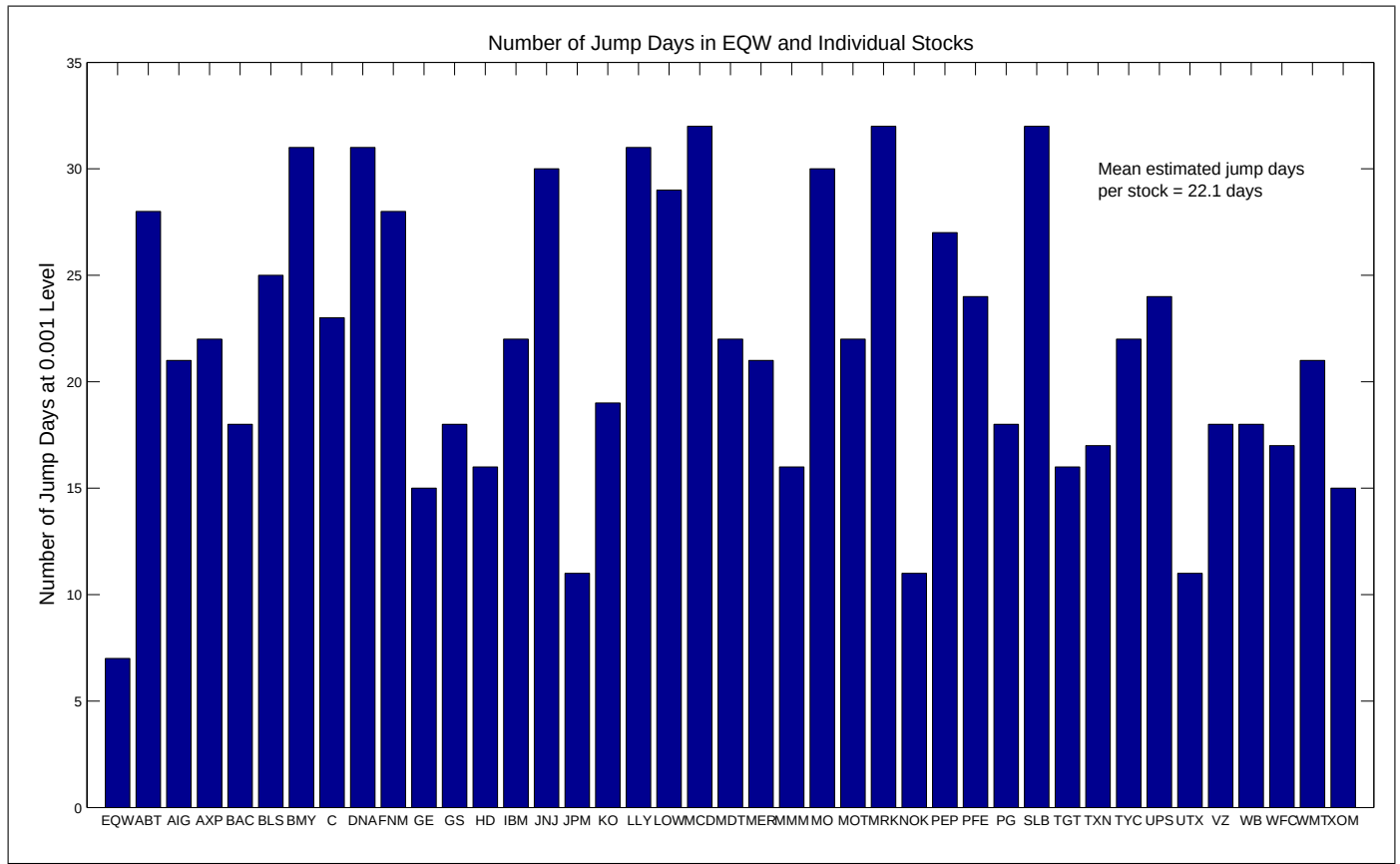


Figure 5: The number of flagged days for EQW and its components from 2001 to 2005.

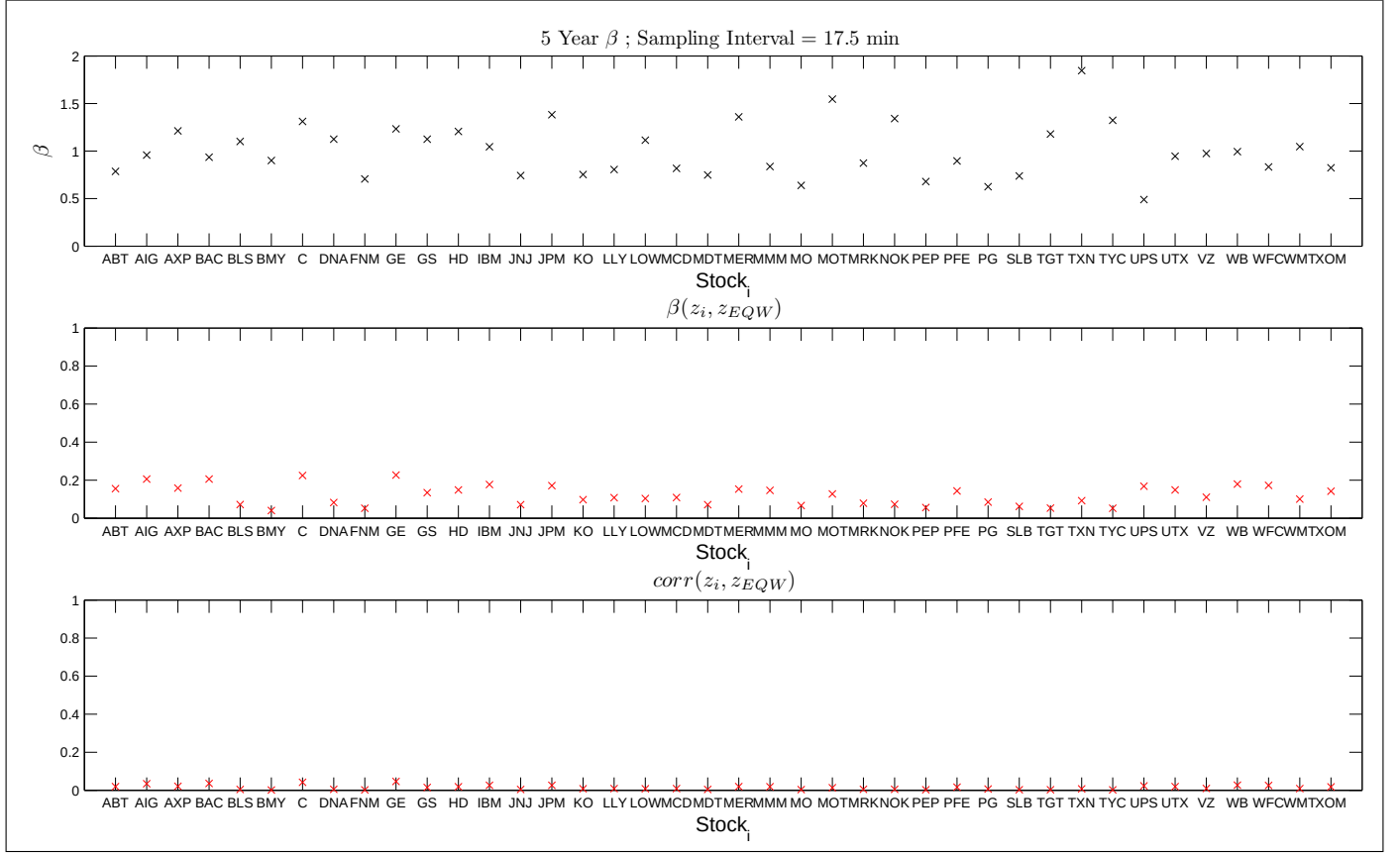


Figure 6: Five-year high frequency β and $corr(z_{EQW,t}, z_{t,i})$

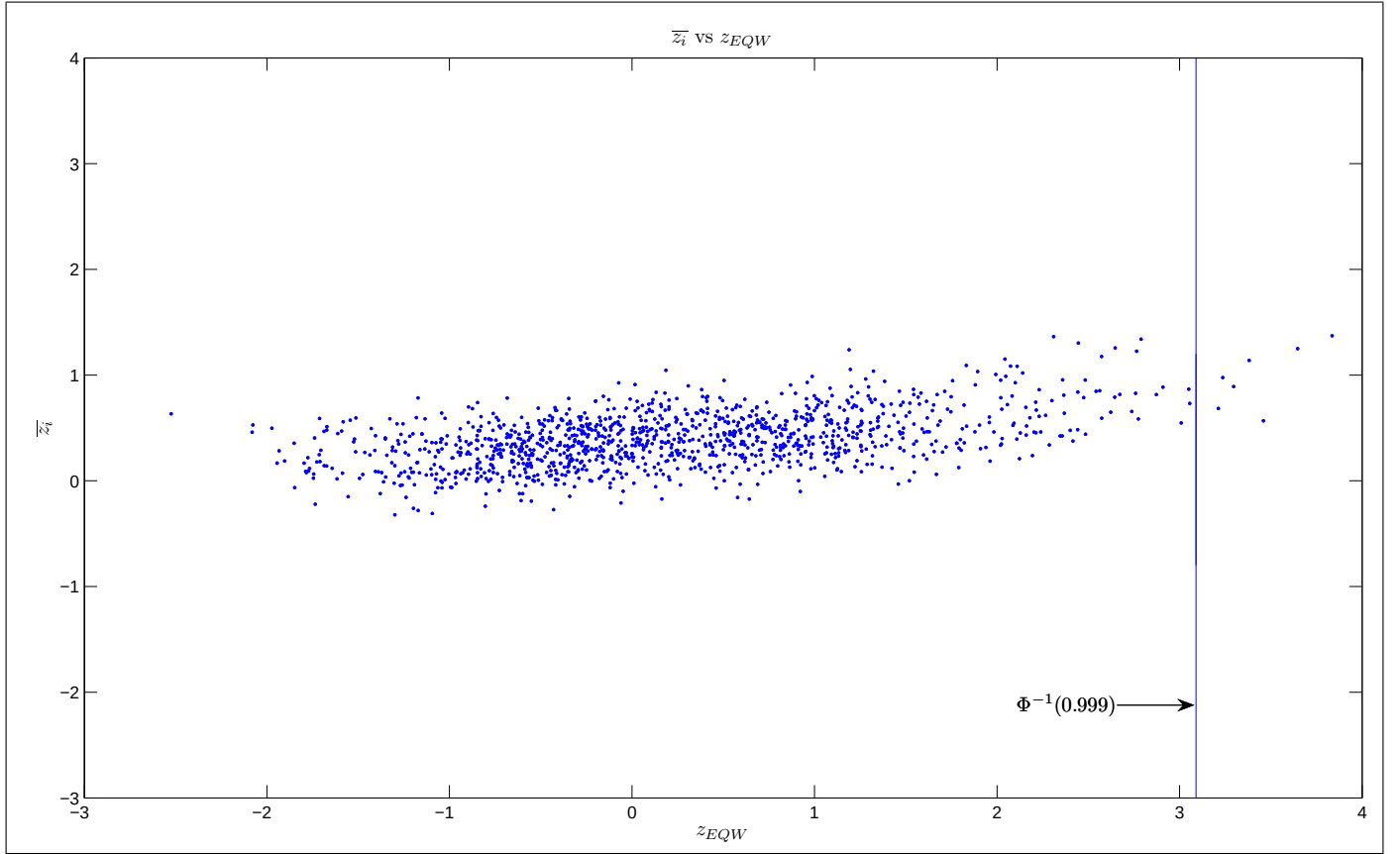


Figure 7: Scatter plot of the average z -statistic of individual stocks and the corresponding z -statistic for EQW.

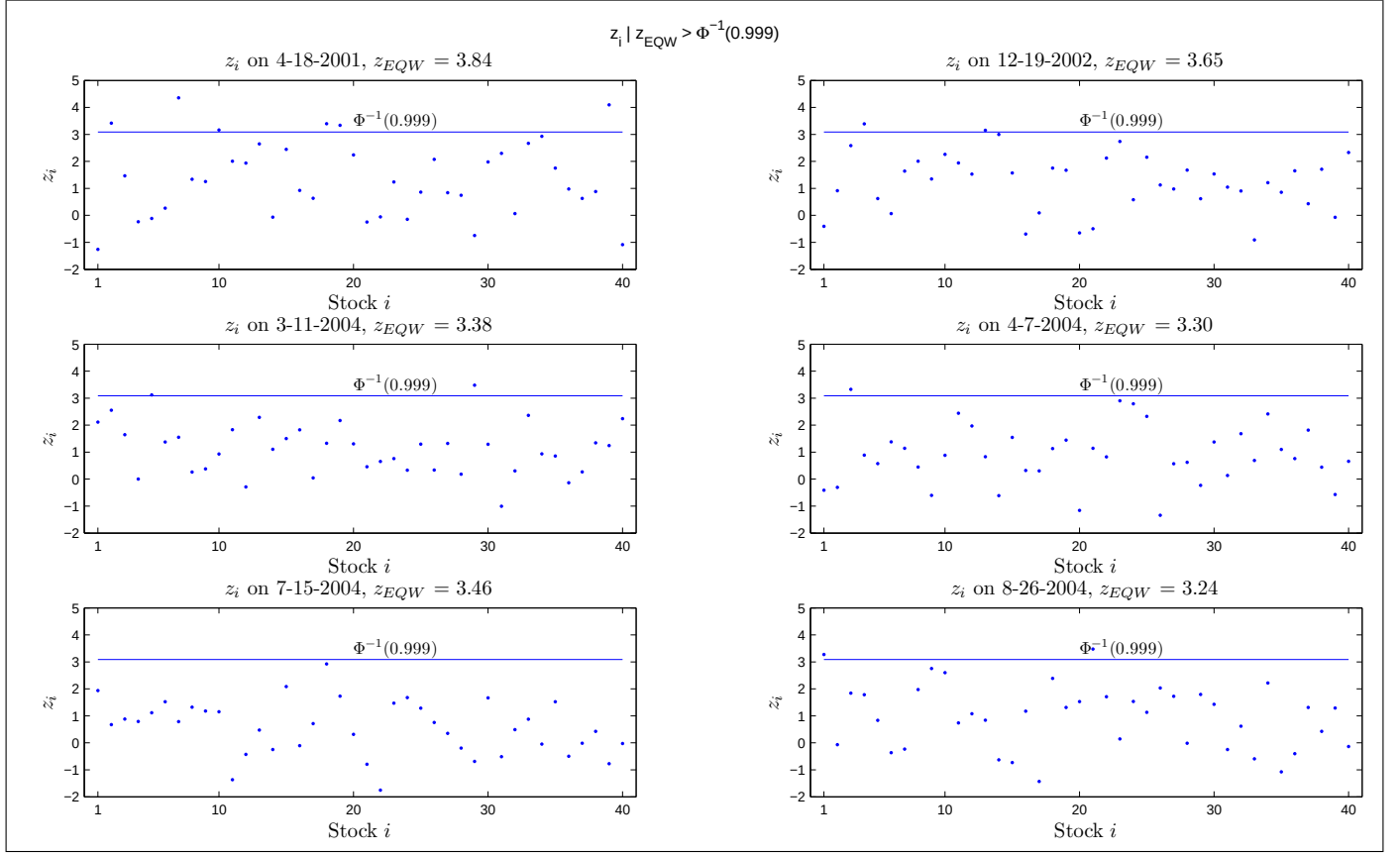


Figure 8: The z -statistics of individual stocks on the six of seven days with the largest z -statistic where EQW is flagged as a jump day

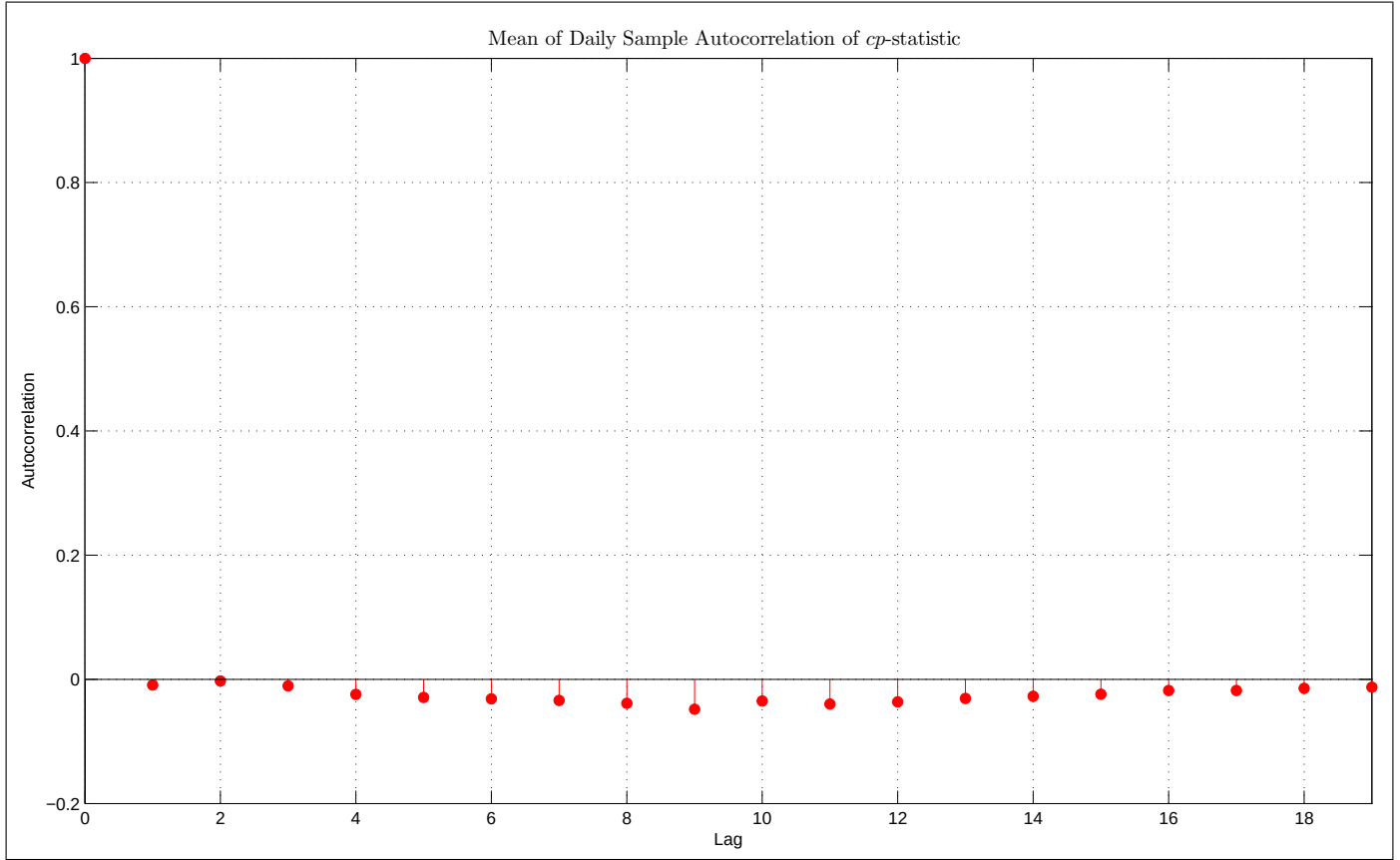


Figure 9: Average serial autocorrelation of the cp -statistics.

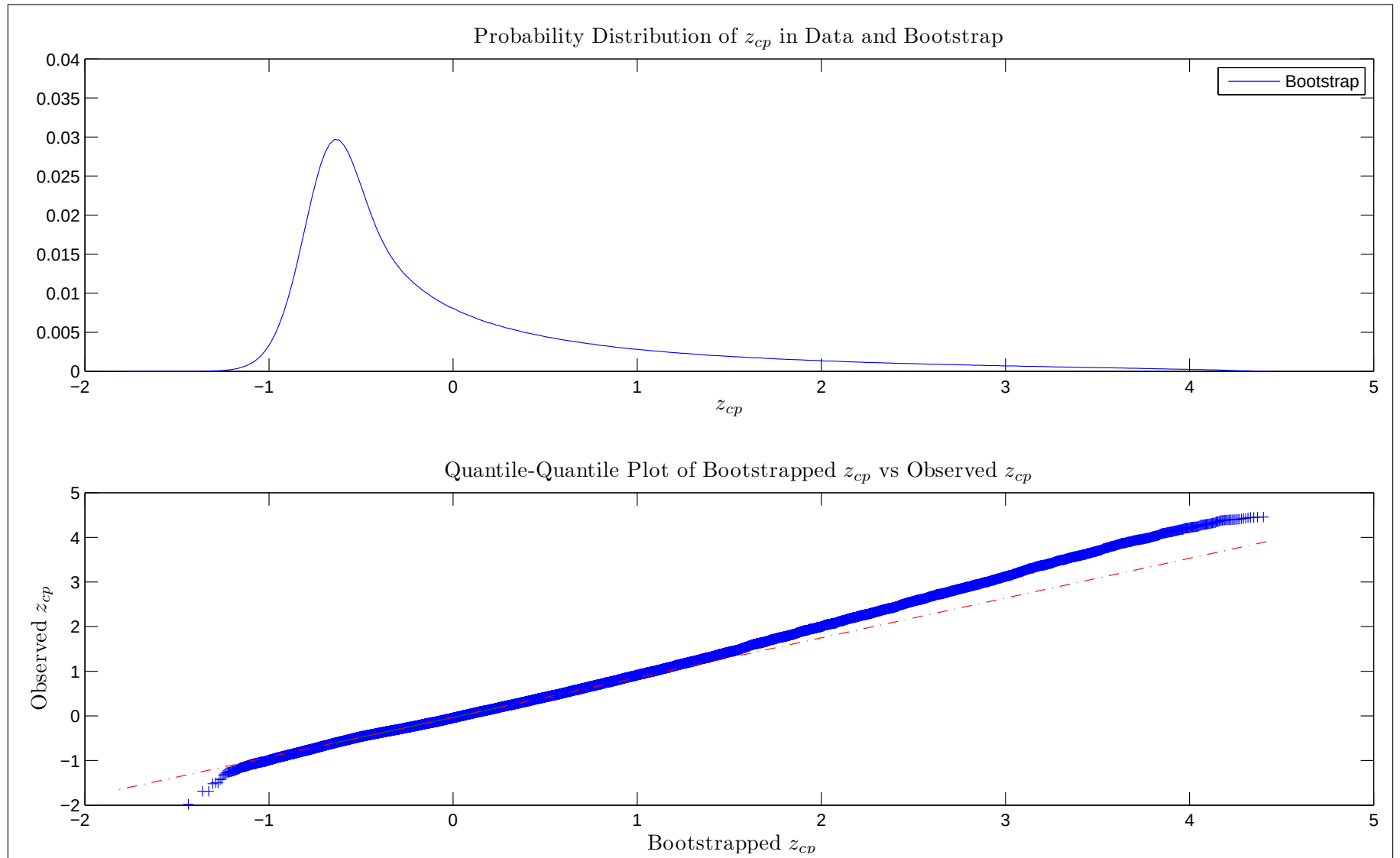


Figure 10: Top: Data-based bootstrap distribution of the z_{cp} -statistic with $n = 40$ stocks and $M = 22$ within-day returns. Bottom: Q - Q plot of the empirical distribution of the observed z_{cp} -statistics relative to the bootstrap distribution.

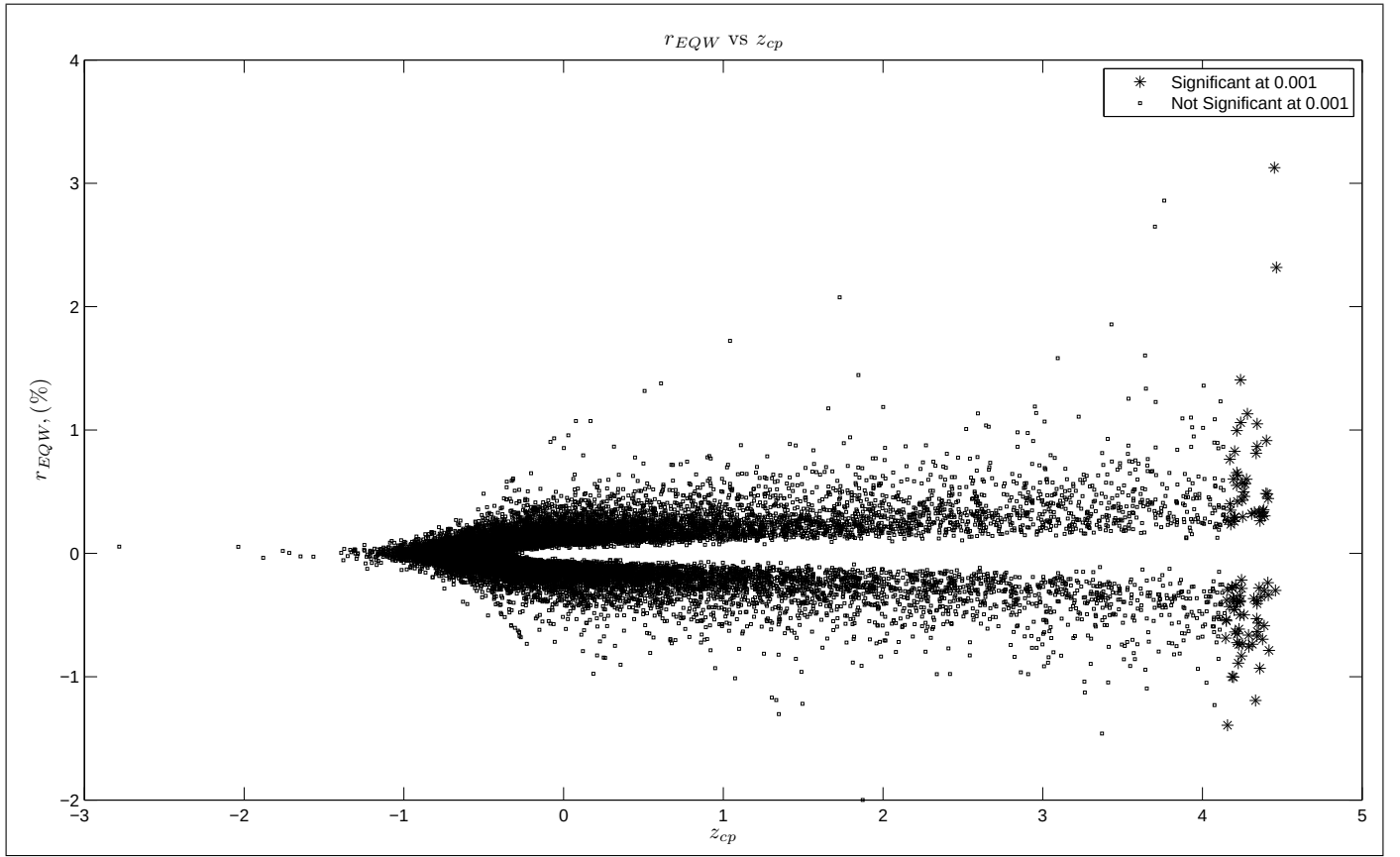


Figure 11: Intra-day EQW returns versus the z_{cp} -statistic.

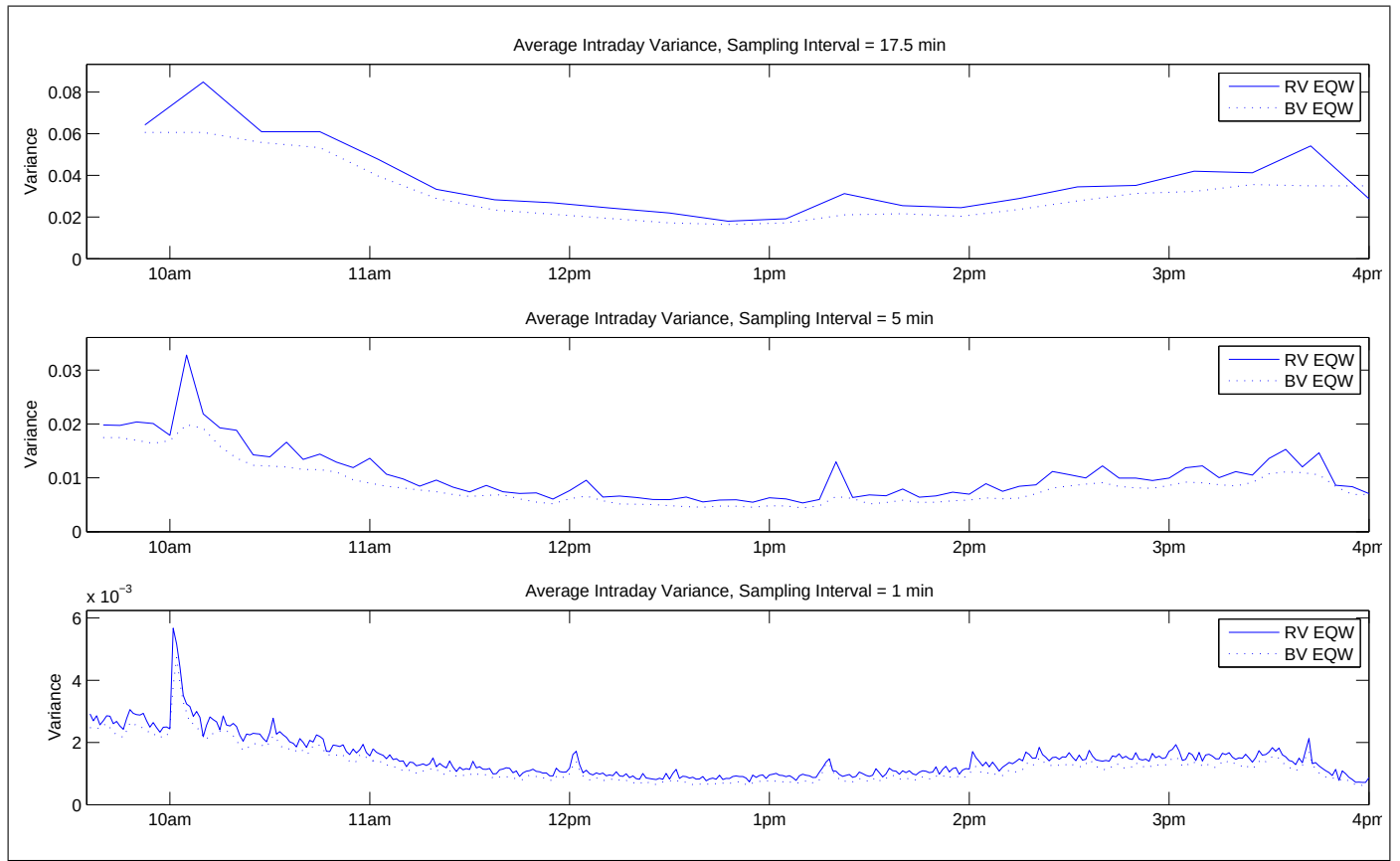


Figure 12: Intraday Realized and Bipower Variation of the EQW Index

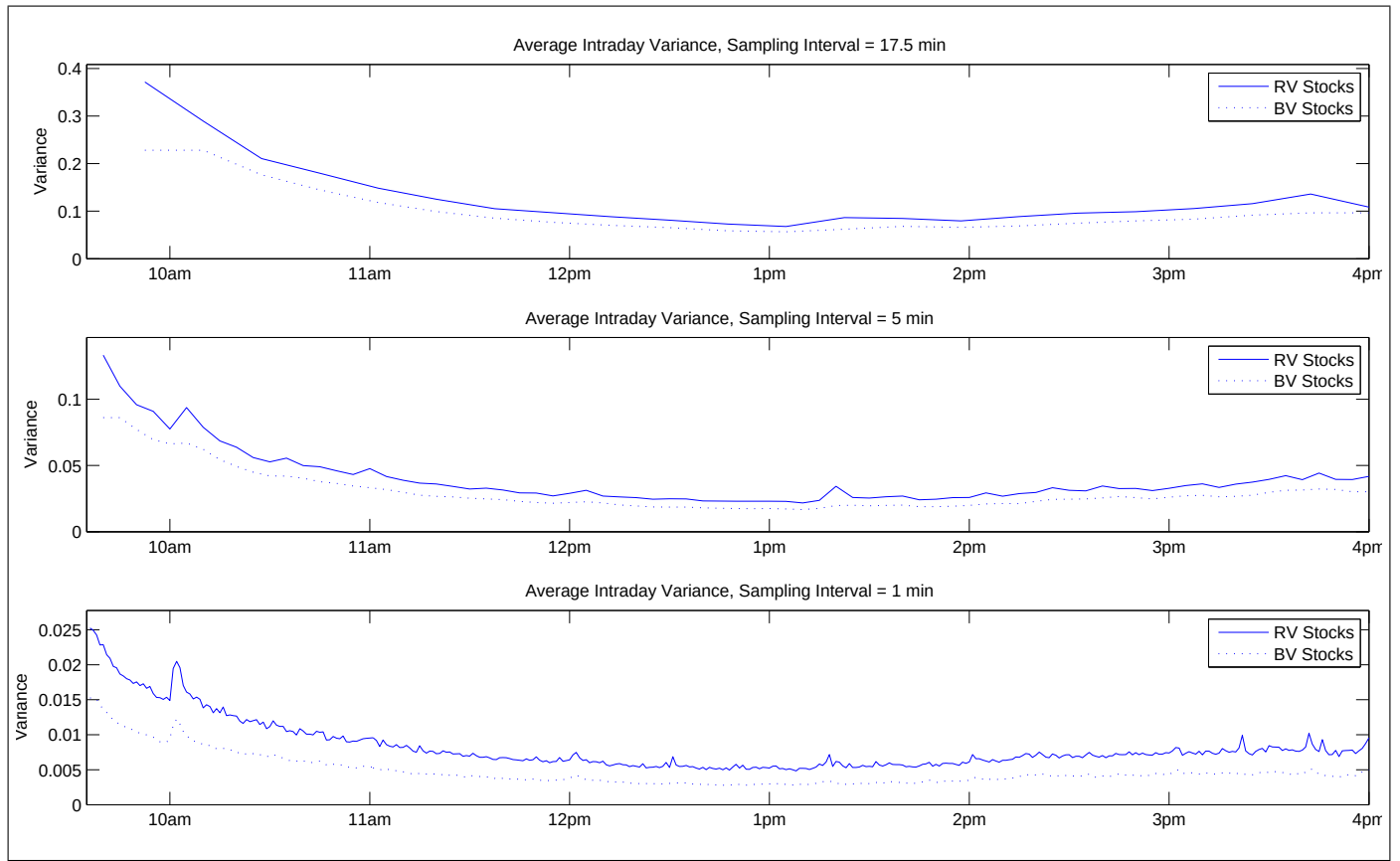


Figure 13: Average Over the 40 Stocks of Intraday Realized and Bipower Variation

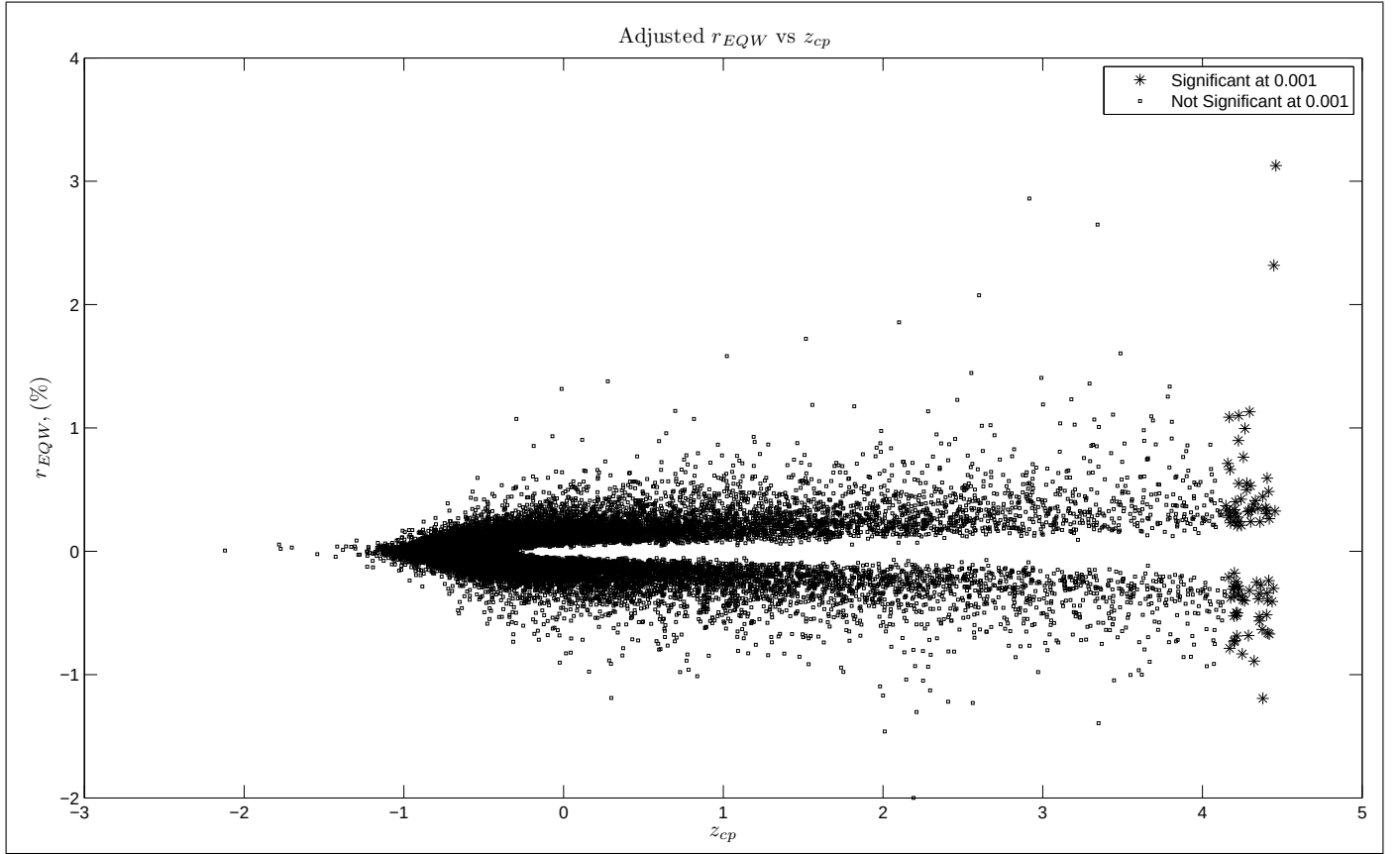


Figure 14: Intra-day EQW Returns versus the z_{cp} -statistic from Adjusted Data.

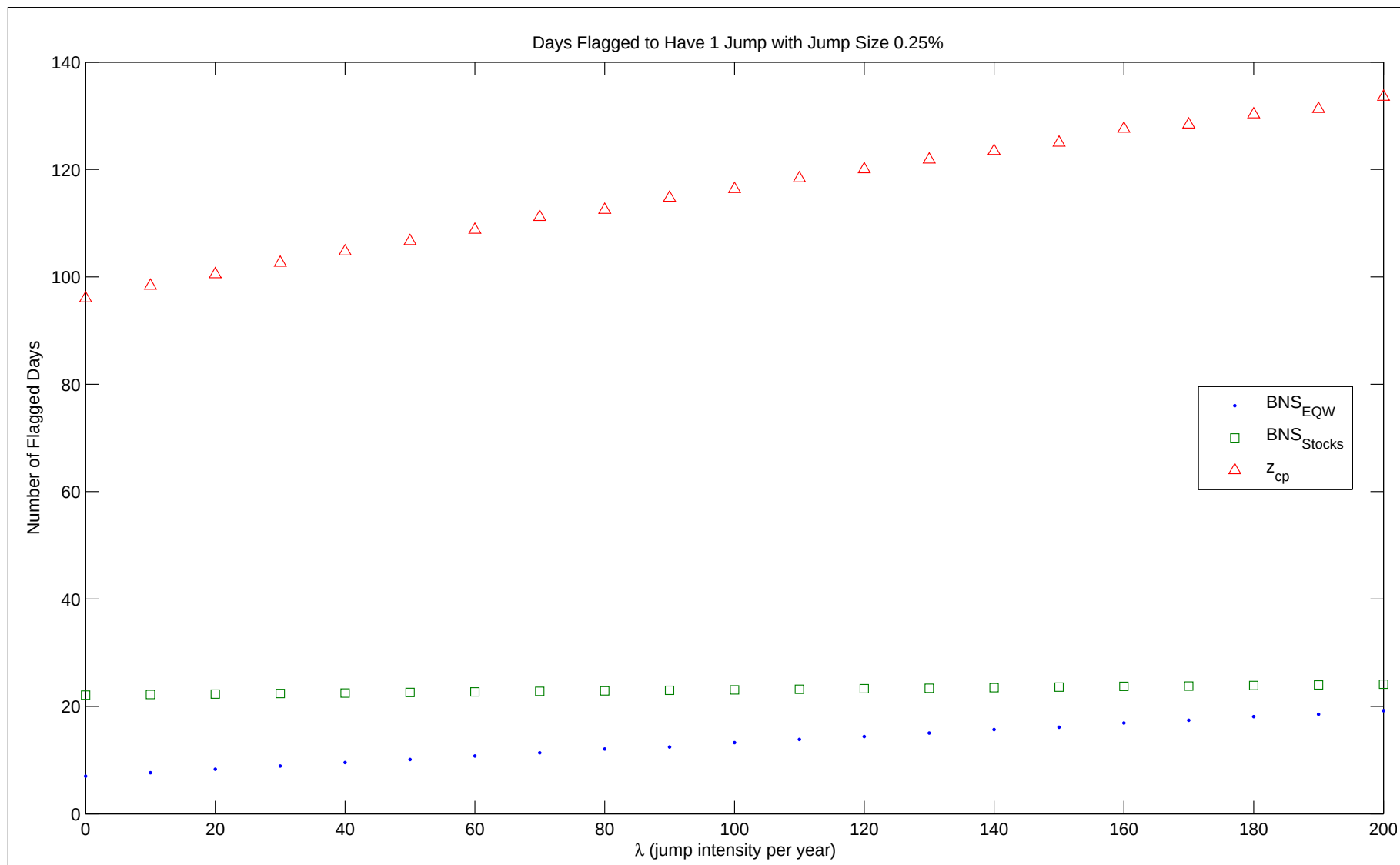


Figure 15:

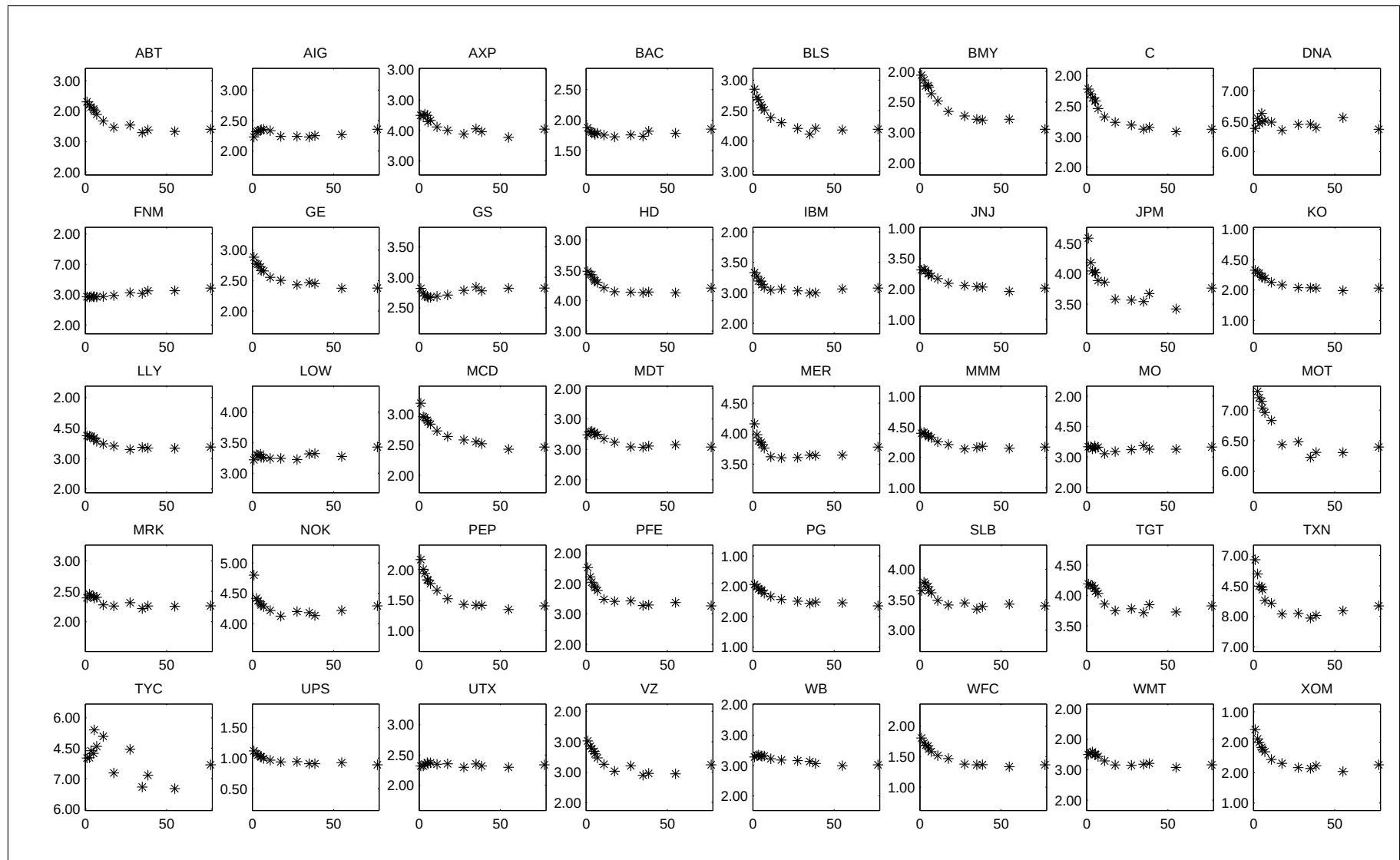


Figure 16: Volatility Signature plots of 40 stocks.

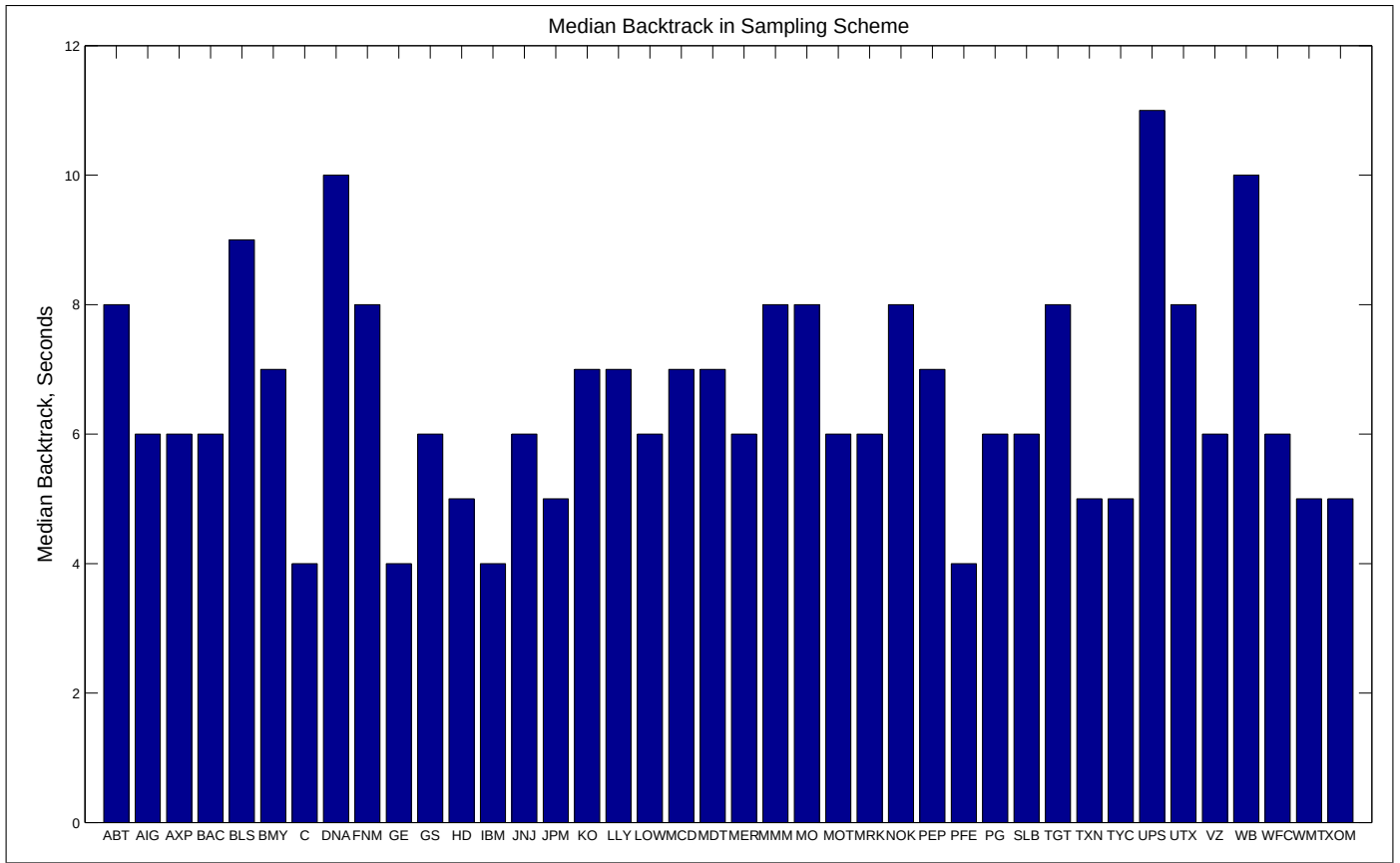


Figure 17: Median backtracking using the previous tick method.

Research Papers 2007



- 2007-6 Charlotte Christiansen: Decomposing European Bond and Equity Volatility
- 2007-7 Stig Møller: Habit persistence: Explaining cross sectional variation in returns and time-varying expected returns
- 2007-8 Charlotte Christiansen, Juanna Schröter Joensen and Jesper Rangvid: Are Economists More Likely to Hold Stocks?
- 2007-9 Thomas Busch, Bent Jesper Christensen and Morten Ørregaard Nielsen: The Role of Implied Volatility in Forecasting Future Realized Volatility and Jumps in Foreign Exchange, Stock, and Bond Markets
- 2007-10 Bent Jesper Christensen, Morten Ørregaard Nielsen and Jie Zhu: Long Memory in Stock Market Volatility and the Volatility-in-Mean Effect: The FIEGARCH-M Model
- 2007-11 Matias D. Cattaneo, Richard K. Crump and Michael Jansson: Optimal Inference for Instrumental Variables Regression with non-Gaussian Errors
- 2007-12 Michael Jansson: Semiparametric Power Envelopes for Tests of the Unit Root Hypothesis
- 2007-13 Jens Perch Nielsen, Carsten Tanggaard and M.C. Jones: Local Linear Density Estimation for Filtered Survival Data, with Bias Correction
- 2007-14 Torben G. Andersen, Tim Bollerslev and Xin Huang: A Reduced Form Framework for Modeling Volatility of Speculative Prices based on Realized Variation Measures
- 2007-15 Viktor Todorov and Tim Bollerslev: Jumps and Betas: A New Framework for Disentangling and Estimating Systematic Risks
- 2007-16 Tim Bollerslev, Michael Gibson and Hao Zhou: Dynamic Estimation of Volatility Risk Premia and Investor Risk Aversion from Option-Implied and Realized Volatilities
- 2007-17 Tim Bollerslev and Hao Zhou: Expected Stock Returns and Variance Risk Premia
- 2007-18 Torben G. Andersen, Tim Bollerslev and Francis X. Diebold: Roughing It Up: Including Jump Components in the Measurement, Modeling and Forecasting of Return Volatility
- 2007-19 Tim Bollerslev, Tzuo Hann Law and George Tauchen: Risk, Jumps, and Diversification