SCHOOL OF ECONOMICS AND MANAGEMENT
FACULTY OF SOCIAL SCIENCES
AARHUS UNIVERSITY

CREATES
Center for Research in Econometric Analysis of Time Series

CREATES Research Paper 2009-8

Jump Testing and the Speed of Market Adjustment

Torben B. Rasmussen

School of Economics and Management
Aarhus University
Bartholins Allé 10, Building 1322, DK-8000 Aarhus C
Denmark

Jump Testing and the Speed of Market Adjustment*

Torben B. Rasmussen†
University of Aarhus and CREATES

February 27, 2009

Abstract

Asymptotic properties of jump tests rely on the property that any jump occurs within a single time interval no matter what the observation frequency is. Market microstructure effects in relation to news-induced revaluation of the underlying variable is likely to make this an unrealistic assumption for high-frequency transaction data. To capture these microstructure effects, this paper suggests a model in which market prices adjust gradually to jumps in the underlying efficient price. A case study illustrates the empirical relevance of the model, and the performance of different jump tests is investigated here and in a simulation study. Evidence indicates that tests based on the largest of scaled price increments perform better than tests comparing measures of variability. Resolving the matter by testing at lower frequencies turns out to be less straightforward.

Keywords: jumps, hypothesis tests, market microstructure noise, high-frequency data.

JEL Classifications: C12, C14, G12

*I would like to thank the Bendheim Center for Finance, Princeton University for its hospitality while I was staying there as a visiting PhD student. The initial work on the paper was started during this stay. I thank Bent Jesper Christensen for comments and discussions, as well as participants at a CREATES seminar for comments on an early version of the paper. CREATES (Center for Research in Econometric Analysis of Time Series) is funded by the Danish National Research Foundation.

†E-mail: trasmussen@econ.au.dk.

Introduction

This paper argues that otherwise discontinuous changes in an economic variable may easily be disguised by market microstructure effects in high frequency observations of the variable, such that jumps will be difficult to detect using recently developed tests.

Discontinuous movement in financial variables have important implications for central issues such as risk management and derivatives pricing. If the variable of interest occasionally moves in jumps, perfect hedges and risk-less replication will only be feasible in exceptional cases, since some risk no longer can be removed by trading frequently enough as in purely continuous models, see Merton (1976) and Naik & Lee (1990). The relevance of including jumps in models of asset prices and interest rates has among others been studied by Bakshi et al. (1997) and Johannes (2004). Deciding whether jumps are needed in the underlying continuous-time model is not a trivial task, since we only have discrete-time observations. Different inference techniques for determining whether the observed movement of a variable is sufficiently described by continuous behavior alone or should include a jump component have therefore appeared in recent literature, see Barndorff-Nielsen & Shephard (2006), Mancini (2006), Aït-Sahalia & Jacod (2009), Jiang & Oomen (2008), and Lee & Mykland (2008).

Inference methods to detect jumps rely on the property that at any discrete observation frequency, no matter how high, a jump will always be included in a single increment of the observed variable. As continuous variation, on the other hand, decreases with the length of the interval, jumps become asymptotically identifiable as the observation frequency is increased. The recent advent of high-frequency data with observations on all transactions or quotations therefore potentially improves empirical implementation of the different testing methods.

Dealing with high-frequency observations is complicated by the possibility that variables may not be perfectly observed. If observations include additive white noise, and this is neglected in the model and estimation method, then the resulting error committed will typically be negligible for observations at low frequency. However, with the amount of noise in an increment not depending on interval length, the signal-to-noise ratio decreases as the frequency of observations is increased. Therefore, the desirability of using as frequent observations as possible is in this case countered by noise becoming an increasing part of observed movements.

Market microstructure, the fine structure of how prices are formed in markets, may affect prices and lead to deviations from the price described by asset pricing theory, the efficient price. For instance, white noise may arise in transactions prices as a consequence of a bid-ask spread set by a market maker to cover operating costs of maintaining a liquid market, cf. Roll (1984). This type of noise and its potential of biasing tests at high frequencies is recognized in the jump testing literature. For example, Aït-Sahalia & Jacod (2009) discuss it in relation to their test, while Jiang & Oomen (2008) and Podolskij & Ziggel (2008) suggest ways to obtain robustness to such noise.

Microstructure effects leading to more complicated noise properties than white noise have received limited attention in relation to testing for jumps. This is in contrast to the situation in the microstructure literature, in which the trading process generally is not seen as independent from changes in the underlying price, thus opening the potential for price related noise. Classical examples of models with microstructure effects related to underlying price movements are the asymmetric information models by Glosten & Milgrom (1985) and Easley & O'Hara (1992), in which buy and sell orders convey private information about asset value and therefore affect the market price permanently.

More relevant in relation to jump tests, though, are microstructure effects that limit the adjustment speed of market prices to changes in the efficient price. Such effects may be related to the market clearing process, e.g. stale limit orders and requirements on the market maker to maintain price continuity. They may be related to agent repositioning, such as time to revaluate the asset, to determine the desired position, and to update orders, or they may be related to dissemination of information, which need not be immediately known by all agents. Some of these effects are discussed by Goldman & Beja (1979), who suggest a model to describe the movement of market prices when such effects are present, and by Hasbrouck (1991) as effects that may cause the trade innovation not to be entirely due to private information. Hasbrouck & Ho (1987) use a model that combines gradual price adjustment with white noise to match the observed autocorrelation pattern in transaction prices, which is negative at first lag and positive for a couple of lags above one. Amihud & Mendelson (1987) investigate differences between clearing house and dealership market types, in relation to a model with gradual price adjustment.

If the speed of market adjustment to changes in the efficient price is limited, then a jump in this underlying price is not reflected in only a single transaction price increment. Since this property is central to jump tests, such effects are likely to cause problems for the tests.

To investigate the impact of effects that limit market adjustment speed on jump tests, I suggest a model inspired by Goldman & Beja (1979), but where gradual adjustment in market prices occurs only in relation to discontinuous movements in the underlying price. Such jumps are mainly due to new information and this model therefore emphasizes effects related to information dissemination and agent repositioning. The behavior of the observed log price, Y, follows a model that is a modification of a standard jump-diffusion model

$$\begin{aligned} dY_t &= \sigma_t dW_t + dJ_t, \\ dJ_t &= \kappa \left(J_t^* - J_t\right) dt + \xi_t d\tilde{W}_t, \qquad J_0 = J_0^*. \end{aligned}$$

The drift term in Y is omitted as the model is applied to short time periods for which this is less relevant, whereas the diffusion term is standard with a stochastic volatility process σ. The jump term is where the model differs from the usual jump-diffusion model. J^* is the compound Poisson process that captures discontinuous changes in the underlying price, but at high-frequency these are only observed in a gradual fashion described by J that enters the log-price Y. As shown by the second equation the observed part of the jump moves toward the underlying jump term with an average speed proportional to the distance of these terms and with some disturbances governed by ξ. The coefficient κ determines the speed of price adjustment and can be seen as a measure of market efficiency. Although the high efficiency of asset markets implies a high value for κ, the suggested model may still produce significant deviations from a standard jump-diffusion model in high-frequency observations.

Jumps in the underlying price enter observed prices only through a jump in the drift term, and the suggested model for market prices therefore remains continuous even when the underlying jump term is active. It is then relevant to consider whether this model should be seen as an alternative for which we want to reject the null model of a standard continuous diffusion. The large κ implied by efficient markets leads to a large, short-lived drift term following an underlying jump, and as a consequence to price behavior that is very different from that in a diffusion model with conventional drift term. The practical implication of accepting that a variable can be described by a continuous model is that one then approximately can trade at every intermediate price when it changes from one level to another, but this may seem as an unreasonable approximation for the suggested model. Thus when the underlying jump term is active, we would most likely want to conclude that this is indeed an alternative model, although it per se is continuous.

A case study illustrates that stock price behavior well described by the suggested model occurs empirically. In the case, transaction prices appear at first glance to jump, but going to the highest frequency, the move turns out to be a series of small changes in the same direction in line with the gradual price adjustment model.

Three jump tests are investigated in relation to gradual price adjustment. The first two compare different estimates of variability of the underlying process over the sample. Thus, Barndorff-Nielsen & Shephard (2006) consider bipower variation relative to quadratic variation, while Aït-Sahalia & Jacod (2009) compare variability at different time scales. Central to these tests is the comparison of consecutive increments in the sample. The third test is inspired by Lee & Mykland (2008) and uses extreme value theory to test if an increment is too big relative to local variation. This test depends less directly on the relation between neighboring increments than the other two tests. In this paper's application the Lee & Mykland (2008) test is modified slightly such that it is directly applicable as a test for jumps in the full series of observations, and small sample extreme value theory, rather than asymptotic, is used to obtain critical values.

The performance of the three jump tests is illustrated in the empirical transaction price data and investigated further in a simulation study, both under standard specifications for the price and with the suggested model. Tests are applied to increments that are sampled from the full record of transactions at a given time interval, δ, starting from an initial point, t_0. The variation in test results over these parameters in the empirical case illustrates the issue of whether a jump can be observed at any short time interval. The simulation study confirms the tests' problems of detecting a jump in the underlying price when the observed price is affected by the microstructure effects in question. Tests on comparisons of consecutive increments are most affected by the noise, while the test using the largest of scaled increments has higher power against this alternative to continuity. It would still be relevant, though, to develop a test methodology aggregating results over starting points.

The next section sets up the testing framework. Section 2 describes tests that compare variability measures, while tests based on the size of the largest increments are discussed in section 3. The empirical case is presented in section 4. Microstructure noise is discussed in section 5, leading to the proposed model, and a re-evaluation of empirical results in this light. The Monte Carlo simulation study is in section 6, while section 7 concludes.

1 Setting

The goal of jump tests is to determine if a continuous model can adequately describe discrete-time observations, or whether it is necessary to include a jump component in the model. The continuous-time price behavior is never fully observed, so the goal of the tests cannot be to decide about continuity objectively. Indeed, sufficient fine-tuning of model parameters can produce sudden changes in continuous models to the same degree as jumps over discrete observation intervals. Therefore, any test to infer from discrete observations about continuity must impose restrictions on the possible processes in the continuous model. A continuous model implies the ability to trade at any intermediate price when the price moves from one level to another, which for example have implications for hedging possibilities. Thus also for practical reasons it makes sense to impose restrictions on the possible continuous model, such that inability to reject continuity is not caused by the possibility that a diffusion model with wild parameter behavior may be able to explain the discrete observations.

For the alternative to continuity, I only consider finite jump activity, although some of the discussed tests have been shown to have power also against infinite active jump terms. This focus is relevant in relation to microstructure effects that may limit market adjustment to news-induced price changes.

The jump tests are presented in the following framework. Let the process Y_t be driven by Brownian Motion, W_t, and a compound Poisson process, J_t, such that

$$Y_T = Y_0 + \int_0^T a_t dt + \int_0^T \sigma_t dW_t + J_T. \tag{1}$$

Here a and σ are progressively measurable processes guaranteeing that (1) has a unique strong solution, which is adapted and right continuous with left limits. The compound Poisson process can be written as $J_T = \sum_{j=1}^{N_T} c_j$, where N_t is a Poisson process with intensity $\lambda_t < \infty$, and c_j are the jump sizes, which are nonzero random variables.

Over a time period $[0, T]$, we observe Y_t at evenly spaced intervals of length δ. This gives a record of observations, $\{Y_{t_0}, Y_{t_1}, ..., Y_{t_n}\}$, where $t_0 \in [0, \delta)$, $t_i - t_{i-1} = \delta$ and $n = \lfloor (T - t_0) / \delta \rfloor$. The main quantities to be analyzed are the observed changes in Y, therefore write the i'th observed increment over an interval of length δ as $y_i^\delta = Y_{t_i} - Y_{t_i - \delta}$. Evenly spaced observations is clearly a simplifying assumption in many applications, e.g. such as for a stock price with observations on either trades or quotes, which in reality occur at random, irregularly spaced times. Therefore, to get evenly spaced observations, a calendar time conversion is necessary from the full data set. This will be discussed when used in the empirical section.

To restrict the possible continuous models, I follow Lee & Mykland (2008) and assume that the drift and volatility functions do not vary too quickly over time,

$$\sup_i \sup_{s \leq \delta} |a_{t_i + s} - a_{t_i}| = O\left(\delta^{1/2 - \varepsilon}\right) \tag{2}$$

$$\sup_i \sup_{s \leq \delta} |\sigma_{t_i + s} - \sigma_{t_i}| = O\left(\delta^{1/2 - \varepsilon}\right), \tag{3}$$

for some $\varepsilon > 0$[1]. These assumptions are also sufficient for the tests by Barndorff-Nielsen & Shephard (2006) and Aït-Sahalia & Jacod (2009), although these authors allow for more flexibility.

The restriction on the volatility function ensures that any large change in discretely observed prices may not equally well be the result of a sudden large increase in volatility and thus come from continuous movement rather than from a jump. The set-up allows for the leverage effect, since nothing constrains price and volatility from being negatively correlated, but as volatility cannot change too quickly, a leverage effect by opposite simultaneous jumps in price and volatility is excluded. It may already now be noted that the price adjustment model hinted at in the introduction violates the assumptions (2)-(3) as the drift includes a jump, and thus presents a different alternative than jumps to a nicely behaved continuous model. The power of the tests against this non-standard alternative will be investigated later. First the jump tests are introduced to test against the standard alternative hypothesis in which jumps are perfectly observed.

It is standard to define the null and alternative hypotheses for the jump tests as

$$H_0 : N_T = 0 \qquad vs. \qquad H_A : N_T > 0. \tag{4}$$

[1]Note, the inclusion of ε allows for functions behaving like $\delta^{1/2} \log(1/\delta)$ as $\delta \to 0$, while still not including functions with changes decreasing at the slower order δ^α for $\alpha > 1/2$. As seen from (12) this matters for letting drift and volatility themselves be driven by Brownian Motion.

Thus, the tests are for whether the realized path includes jumps, not for whether a jump term is present in Y. This way of defining the hypotheses for jump tests differ from conventional testing methodology, since it is not a test of a property of the hypothesized population from which data is a realization. In theory we should test if a jump component, J, is included in Y or not, but a continuous model can never be distinguished from one with a jump term that has not jumped yet, $N_T = 0$. Thus the best we can do is to test for jumps up to T, though finding $N_T = 0$ does not preclude the presence of a so far inactive jump term. The next two sections present the jump tests.

2 Tests Comparing Measures of Variability

2.1 Quadratic and Bipower Variation

While quadratic variation measures variation from both continuous movements and jumps, bipower variation is designed to only capture variation from the continuous part. Therefore, when estimators of these quantities are properly compared, the jump component can be separated out and a test of continuity is obtained.

The quadratic variation of Y_T in (1) is the integral of squared volatility plus the sum of squared jumps. Realized quadratic variation, the sum of squared increments, is a consistent estimator, as it can be shown to converge in probability to quadratic variation as $\delta \to 0$,

$$[Y_\delta]_T = \sum_{i=1}^{n} \left(y_i^\delta\right)^2 \xrightarrow{P} \int_0^T \sigma_t^2 dt + \sum_{j=1}^{N_T} c_j^2. \tag{5}$$

As an estimator of integrated variance robust to jumps, Barndorff-Nielsen & Shephard (2004) introduced realized bipower variation, the sum of the product of consecutive absolute increments. In a process with finitely many jumps, the probability of jumps in consecutive increments goes to zero as the time interval is decreased. Therefore, possible jump increments get multiplied by neighboring small increments from continuous movement, and in the limit jumps do not affect bipower variation. It is therefore possible to show that realized bipower variation converges in probability to scaled integrated variance,

$$\{Y_\delta\}_T^{[1,1]} = \sum_{i=2}^{n} \left|y_i^\delta\right| \left|y_{i-1}^\delta\right| \xrightarrow{P} \mu_1^2 \int_0^T \sigma_t^2 dt, \tag{6}$$

as $\delta \to 0$. The constant μ_1 is the first absolute moment in the standard normal distribution.

Comparing the two variability estimators, we see that the jump term becomes asymptotically identified. The probability limit under respectively the null and alternative hypotheses for the ratio between the two estimators scaled by μ_1^{-2} can be seen from (5) and (6) to be

$$\hat{R}(\delta) = \frac{\mu_1^{-2}\{Y_\delta\}_T^{[1,1]}}{[Y_\delta]_T} \xrightarrow{P} \begin{cases} 1 & H_0 \\ 1 - \frac{\sum_{j=1}^{N_T} c_j^2}{\int_0^T \sigma_t^2 dt + \sum_{j=1}^{N_T} c_j^2} & H_A. \end{cases} \tag{7}$$

Barndorff-Nielsen & Shephard (2006) obtain results for the asymptotic distribution for both this statistic, as well as for a linear difference statistic, and construct feasible tests. Since the adjusted ratio test seems to perform best in their Monte Carlo study, I will focus on this test.

Appropriately scaled, the ratio statistic converges in distribution to a $N(0,1)$ variable under the null of no jumps. From (7) small values are critical to the null and a one-sided approximate confidence interval can be constructed. At significance level α the critical value for the ratio statistic is

$$C_R^\alpha = 1 - z_{1-\alpha}\left(\delta\vartheta \max\left\{T^{-1}, \frac{\{Y_\delta\}_T^{[1,1,1,1]}}{\left(\{Y_\delta\}_T^{[1,1]}\right)^2}\right\}\right)^{\frac{1}{2}}.$$

Here $\vartheta = \left(\pi^2/4\right) + \pi - 5$, $z_{1-\alpha}$ is the $1-\alpha$ quantile of the standard normal distribution, and we have to compute the realized quadpower variation, $\{Y_\delta\}_T^{[1,1,1,1]} = \delta^{-1}\sum_{i=4}^{\lfloor T/\delta \rfloor}\prod_{j=0}^{3}\left|y_{i-j}^\delta\right|$.

2.2 Variability at Different Sampling Frequencies

Consider the sum of absolute increments to the p'th power,

$$\hat{B}(p,\delta)_T = \sum_{i=1}^{n}\left|y_i^\delta\right|^p. \tag{8}$$

The test suggested by Aït-Sahalia & Jacod (2009) exploits that the limit for $p > 2$ of the sum in (8), as δ goes to zero, is different with and without jumps in the process. When jumps are present $\hat{B}$ converges to a finite, non-zero limit, whereas if the process is continuous the limit is zero and the convergence speed depends on the sequence of δ's going to zero. The idea is therefore to compare (8) at different interval lengths, δ. If the sums are of same size, the evidence points toward jumps in the process, whereas if they are sufficiently different, it points toward continuity.

$\hat{B}$ is equal to realized quadratic variation for $p = 2$, $\hat{B}(2,\delta)_T = [Y_\delta]_T$, in which case we just saw that both continuous and jump variation are captured by the statistic. On the other hand, when $p > 2$ discontinuous movements will dominate as δ decreases,

$$\hat{B}\left(p,\delta\right)_T \xrightarrow{P} \sum_{j=1}^{N_T} \left|c_j\right|^p \quad for \quad p > 2. \tag{9}$$

When there are no jumps, the limit in (9) is clearly zero, but by scaling appropriately a positive, finite limit is obtained. We have that

$$\frac{\delta^{1-p/2}}{\mu_p}\hat{B}\left(p,\delta\right)_T \xrightarrow{P} \int_0^T \left|\sigma_t\right|^p dt = A\left(p\right)_T \quad under \quad H_0, \tag{10}$$

as $\delta \to 0$, where μ_p is the p'th absolute moment in the standard normal distribution.

Now compare the $\hat{B}$ statistic at two time-scales, δ and $m\delta$, for some integer $m > 1$. Based on the convergence properties in (9) and (10), the ratio of $\hat{B}$ at the slower time-scale to that at the faster satisfies

$$\hat{S}\left(p,m,\delta\right) = \frac{\hat{B}\left(p,m\delta\right)_T}{\hat{B}\left(p,\delta\right)_T} \xrightarrow{P} \begin{cases} \frac{\delta^{1-p/2}}{(m\delta)^{1-p/2}} = m^{p/2-1} & H_0 \\ \\ 1 & H_A, \end{cases} \tag{11}$$

and therefore gives a statistic, $\hat{S}$, that can be used to discriminate between jumps and continuity. The design parameters p and m are set to $p = 4$ and $m = 2$ as suggested by Aït-Sahalia & Jacod (2009).

To determine whether $\hat{S}$ is far enough from the asymptotic value under H_0 to reject this hypothesis, the distribution of the statistic under the null is needed. Aït-Sahalia & Jacod (2009) show that with appropriate scaling $\hat{S}$ converges in distribution to an $N\left(0,1\right)$ variable as $\delta \to 0$ when there are no jumps. For fixed δ, this is used as an approximate distribution for $\hat{S}$, and from (11) small values are critical for the null hypothesis. At α level of significance the critical value can be found to be

$$C_{\hat{S}}^{\alpha} = 2 - z_{1-\alpha}\left(\delta M\left(4,2\right)\frac{\hat{A}\left(8,\delta\right)_T}{\hat{A}\left(4,\delta\right)_T^2}\right)^{1/2},$$

Again, $z_{1-\alpha}$ is the $1-\alpha$ quantile of the standard normal distribution, $M\left(4,2\right) = 204$, and $\hat{A}\left(p,\delta\right)_T = \sum_{i=1}^{n}\left|y_i^\delta\right|^p \mathbf{I}\left\{\left|y_i^\delta\right| \leq \psi\delta^\varpi\right\}$ is the realized truncated p'th variation, which for constants $\psi > 0$ and $\varpi \in \left(0,\frac{1}{2}\right)$ estimates $A\left(p\right)_T$. The choice of ψ and ϖ is important for finite sample properties of the test, since $\hat{A}\left(p,\delta\right)_T$ should only include y_i^δ's that are due to continuous movement. Recommended values are ψ at about $3-5$ times of average σ and ϖ close to $\frac{1}{2}$.

3 Tests Based on Largest Increments

The possible variation due to continuous movement decreases with the length of the time interval. Thus for sufficiently small intervals, jumps can be detected as increments, y_i^δ, that are numerically too large to come from continuous variation. To base a test on this observation, it's necessary to determine how much the process can move continuously, or equivalently, to determine the distribution of the largest increments generated by continuous movement alone. Results about this distribution ultimately come from the independent, normally distributed increments of the Wiener process, but a crucial element is to account appropriately for scaling by the volatility term, which may be time-varying.

Identification as large increments Over short enough time periods, the diffusion term of the process (1) is much larger than the drift term. The first step to finding an upper limit on continuous movement is therefore to determine how much Brownian Motion can move. The modulus of continuity, $w_f(\delta)$, of a function f measures the the largest change in function value over intervals less than or equal to δ. For Brownian Motion this is a random quantity, which can be shown to have the property

$$\frac{\sup |W_t - W_s|}{\sqrt{2\delta \log(1/\delta)}} \to 1 \qquad (a.s.) \qquad for \qquad \delta \to 0, \tag{12}$$

for $0 \leq s < t \leq 1$ and $|t-s| \leq \delta$. Therefore, for short time intervals Brownian movements will approximately be of size $\sqrt{2\delta \log(1/\delta)}$.

Allowing for volatility, when this is bounded, will not change the convergence order of the diffusion term, $O_p\left(\sqrt{\delta \log(1/\delta)}\right)$, and this dominates the drift term. In contrast, given that there is a jump in the interval $(t_i - \delta, t_i]$, the size of the jump movement does not depend on the length of the interval, it is of order $O_p(1)$.

Together, these results confirm that jumps can be observed as increments at sufficiently short time periods that are too large relative to that of Brownian increments times a constant to adjust for volatility. Further results on almost sure identification of jumps by this method are given in Mancini (2004) and Mancini & Reno (2006), who apply this to clean for jumps in truncated quadratic variation estimators of volatility. For inference on whether the process included jumps, we need distributional results.

Approximate normality The independent, normal increments of the Wiener process are perturbed by the general stochastic volatility term in the process (1) to more general distributions for the increments of the diffusion term. For short time intervals and well behaved volatility, though, the normal distribution remains a valid approximation. Define the set of i's for which the process has no jumps in the interval $(t_i - \delta, t_i]$ as G_δ. Then for short time intervals we have approximately that

$$y_i^\delta \overset{a}{\sim} N\left(\int_{t_i-\delta}^{t_i} a_s ds, IV_i^\delta\right) \qquad for \qquad i \in G_\delta,$$

for integrated variance, $IV_i^\delta = \int_{t_i-\delta}^{t_i} \sigma_s^2 ds$. With estimates of integrated drift and variance, the increments, y_i^δ, can be scaled to get a series of approximately standard normal variables. It is then relatively straightforward to obtain results for the largest scaled increments in G_δ by using available results for extremes in the standard normal distribution.

For any reasonable drift term in (1), changes in Y due to this term over short time intervals are much smaller than changes in the diffusion term. Also, relatively large errors in estimation of the drift will decrease the precision in estimates of integrated variance. Therefore, for standardizing observed increments, the drift is set to zero, and the focus will be on adjusting for integrated variance. This follows Lee & Mykland (2008), and standard practice in many applications with short time intervals.

Estimation of Integrated Variance Scaled bipower variation, $\mu_1^{-2}\{Y_\delta\}_T$ was discussed in (6) as a jump robust estimator of integrated variance over the period $[0, T]$. Here, we need the estimate over a short time interval of length δ and cannot rely on observations within the interval. Instead, due to the assumption that volatility doesn't change too quickly, IV_i^δ can be estimated using a local window around the interval of interest. Specifically, let the local window be K increments to either side of the i'th, such that we consider a window of $2K + 1$ increments, $\{y_{i-K}^\delta, ..., y_i^\delta, ..., y_{i+K}^\delta\}$. Then the estimator based on bipower variation is

$$\widehat{IV}_i^\delta = \frac{\mu_1^{-2}}{2K-1} \sum_{j=i-K+1}^{i+K} \left|y_j^\delta\right| \left|y_{j-1}^\delta\right|. \tag{13}$$

For this to be a consistent estimator, for $\delta \to 0$, the window length parameter K must be adjusted to satisfy $K \in O_p\left(\delta^{-\nu}\right)$, for $\nu \in \left(\frac{1}{2}, 1\right)$. This condition ensures that the window length decreases with δ, since $K\delta \to 0$, while the number of increments still increases fast enough.

Scaled Increments With the estimate of integrated variance in place, the increments can be scaled to obtain a series of variables whose distribution is approximately equivalent to a series of independent $N(0, 1)$ variables for non-jump intervals,

$$\hat{z}_i^\delta = y_i^\delta / \sqrt{\widehat{IV}_i^\delta} \overset{a}{\sim} N(0,1) \qquad for \qquad i \in G_\delta. \tag{14}$$

This approximation follows from the asymptotic result in theorem 1 of Lee & Mykland (2008), from which it follows that

$$\sup_i \left| \hat{z}_i^\delta - z_i \right| = O_p\left(\delta^{3/2-\zeta-\nu-\varepsilon}\right) \qquad for \qquad i \in G_\delta, \tag{15}$$

where the z_i's are independent $N(0,1)$ random variables. The constant ν is from the condition on the estimation window for integrated variance, ζ is a constant satisfying $0 < \zeta < 3/2 - \nu$, and (15) holds for all $\varepsilon > 0$. In other words, the condition holds for positive exponents, so the largest distance between $\hat{z}_i^\delta$ and z_i goes to zero with δ, and the approximation error in (14) decreases as time intervals become shorter.

To detect jumps we can, as argued above, use the property that conditional on a jump in $(t_i - \delta, t_i]$, the size of the jump movement does not decrease with δ. Since $IV_i^\delta \to 0$ for $\delta \to 0$, the scaled increments tend to infinity for jump intervals,

$$\left| \hat{z}_i^\delta \right| \to \infty \qquad for \qquad \delta \to 0 \qquad and \qquad i \notin G_\delta. \tag{16}$$

Comparing this to (14) a test for jumps should check if the largest scaled increments are abnormally large relative to extremes in independent draws from the standard normal distribution.

Results from Extreme Value Theory For a series of n independent variables with symmetric distribution around zero, $\{X_1, ..., X_n\}$, define the largest absolute value as $Q(n) = \max_{i=1,...,n} |X_i|$. The probability of no X_i's numerically larger than u is the probability of no successes in n independent draws with probability equal to that of $|X_i| > u$. Hence, the distribution function for the largest absolute value of n independent $N(0,1)$ variables can be calculated as

$$P(Q(n) \leq u) = P\left(V_{n,2(1-F(u))} = 0\right), \tag{17}$$

where $V_{n,p} \sim Bin(n,p)$ and F is the cdf. for each X_i, see the appendix for details.

Instead of the exact extreme value result in (17), Lee & Mykland (2008) use asymptotic theory. Properly scaled, extremes of samples from the normal distribution converge in distribution to the Gumbel distribution, also called the double exponential distribution, which has cdf. $\Lambda(x) = \exp(-\exp(-x))$. Using this asymptotic result to approximate the distribution in finite samples gives the distribution function in closed form and therefore simplifies calculation of critical values. On the other hand, the convergence rate of normal extremes to the Gumbel distribution is very slow. Even for optimal choices of normalizing constants, the rate of convergence will not exceed $(\log n)^{-1}$. This is slower than the power convergence obtained for $\hat{z}_i^\delta$ to standard normal variables in (15). So despite the relatively large number of increments obtained from high-frequency data, using the Gumbel distribution will be a further approximation. Though (17) is not in closed form, evaluation and inversion is straightforward on standard numerical software, which must be applied anyway to handle the large high-frequency data set.

3.1 Extreme Value Test

The results in the previous subsections are now collected to a test for jumps in the path of Y up to T. Under the null hypothesis of no jumps, all $i \in G_\delta$, so from (15) the sequence $\hat{z}_1^\delta, ..., \hat{z}_n^\delta$ is approximately independent, standard normally distributed. Therefore, the largest absolute value in this sequence,

$$\hat{Q}(\delta) = \max_{i=1,\ldots,n} \left|\hat{z}_i^\delta\right|, \tag{18}$$

has a distribution approximately equal to that of $Q(n)$ in (17) with $F = \Phi$. Note that δ is used to indicate the length of the sequence, since for a given time period this implicitly determines the number of observations, $n = \lfloor (T - t_0)/\delta \rfloor$. Under the alternative hypothesis some interval has a jump, $i \notin G_\delta$, and then from (16), the $\hat{Q}$ statistic will be large relative to the distribution under the null hypothesis. The extreme value test (EV test) for jumps is then a one-sided test, for which the critical value for $\hat{Q}(\delta)$ at α level can be found from

$$C_Q^\alpha = \{u : P(Q(n) \leq u) = 1 - \alpha\}.$$

This is solved numerically, using (17), for example for $\alpha = 5\%$ and $n = 760$, corresponding to a trading day with observations every 30 seconds, the critical value would be $C_Q^\alpha = 2.88$.

The test in Lee & Mykland (2008) is, in contrast to the formulation here, a test for jumps in each individual time interval. Their test thus compares each $\left|\hat{z}_i^\delta\right|$ to the distribution of the maximal absolute increment over the full period. By this method the test in each interval will be conservative, i.e. have lower probability of spuriously detecting a jump than the chosen significance level. The full sample path can be tested for jumps by applying the individual test to all increments, but as seen in their table 4, the resulting size is much lower than the chosen one. By directly focusing on a global test, the method in the present paper implies that the size of the global test is approximately the one chosen for the critical value, as will be confirmed in the simulation study. This makes the comparison to other jump tests more even.

4 Empirical Case

The performance of the jump tests is now studied in relation to a case of intraday transaction price observations for a single stock on the New York Stock Exchange. This series was chosen to investigate the performance of tests in a case where the price appears to change abruptly following a company news announcement.

4.1 GM October 6, 2006

Data Tick-by-tick data for General Motors (GM) on October 6, 2006 is collected from the NYSE trades and quotes (TAQ) database. This gives the full record of intraday transactions, including price, volume, and time stamp for each trade. In total 128,702 trades were observed during this day.

I refer to Brownlees & Gallo (2006) for a discussion of data handling issues in dealing with this type of ultra high-frequency data. First, wrong or inaccurate ticks are removed as indicated by the CORR and COND fields in the TAQ database. This removes 376 trades. Next, a filter is applied to remove records that do not seem to come from plausible market activity. These are trades with prices significantly different from the prices of the surrounding trades. I use the filter suggested by Brownlees & Gallo (2006) to identify these trades: For each observation i, calculate the d-trimmed mean and standard deviation, $s_i(k,d)$, in a window of $k/2$ observations to either side. Then remove observation i if the price is further away from the trimmed mean than $3s_i(k,d)+\phi$, where ϕ is a granularity parameter set to avoid very small thresholds due to a sequence of trades at the same price. With parameter values at $k=60$, $d=10\%$, and $\phi=.02$, the filter removes 432 observations, leaving a cleaned data set of 127,894 observations.

The time stamp of each transaction is recorded in seconds, so during the 6.5 hour trading day, corresponding to 23,400 seconds, several trades have the same time stamp. For seconds with multiple trades, I choose to use the median of the recorded prices, and when an even number of different prices are observed, volume decides which of the middle prices to use. This reduces the data set to 15,060 observations with different time stamps, which will be considered as the basic, cleaned data set. Figure 1 illustrates the filtering procedure over an interval of four minutes with high activity. The figure shows all observations, with those removed by the filter marked by a circle, and the final series with a single price for each second that has an observation shown as a line.

The top panel in figure 2 shows the cleaned price process over the full day, clearly showing that the abrupt price drop at around 12^{PM} stands out. The accumulated volume is shown in the bottom panel of figure 2. Notice the increase in speed of trade following the price drop. New information about the company was released at about 12^{PM} that day, when it was announced that a central member of the board had chosen to resign, citing concerns over the company's ability to compete in the market among other reasons for the resignation.

Tests To avoid adjusting tests to account for unevenly spaced observations, the price series will be subsampled at intervals of length δ to get data in proper format. This calendar time conversion also facilitates the study of how the tests perform at different observation frequencies as δ is easily adjusted. Thus, for starting point t_0, the subsampling procedure selects observations at seconds $\{t_0, t_0 + \delta, t_0 + 2\delta, ...\}$. If for some $t_0 + i\delta$ there is no observation, the most recently recorded transaction prior to this time is used, similarly to the method in Andersen, Bollerslev, Diebold & Ebens (2001). Since the GM data set has observations at about 2/3 of the seconds during the trading day and the shortest sampling interval used is 15 seconds, the approximation implied by the procedure seems acceptable.

It is important to check results over different sample starting points, t_0. For a time step of δ seconds, there are δ possible starting points, $\{0, 1, ..., \delta - 1\}$, that give rise to completely different sets of increments. For the BNS and EV tests, this is also the number of starting points to consider, while for the AJ test, one also samples at intervals of length 2δ, and thus must check twice as many starting points.

To estimate integrated variance, $\widehat{IV}_i^{\delta}$ in (13), for the EV test, the window size parameter is set to $K = \left\lceil 120\sqrt{30/\delta} \right\rceil$, where δ is measured in seconds. This implies, e.g., that a local window of 2 hours is used for δ at 30 seconds. For t_i with less than K observations to either side, the window is shifted to keep the same length, while for the most infrequent sampling method, δ at 10 minutes, a constant IV estimate over the full day is used.

The AJ, BNS, and EV jump test statistics, $\hat{S}(4, 2, \delta)$, $\hat{R}(\delta)$, and $\hat{Q}(\delta)$, were calculated for intervals of length 15 and 30 seconds, and 1, 2, 5, and 10 minutes. Due to the large number of tests considered, the results are shown graphically in figures 3 and 4 together with critical values for significance level $\alpha = 5\%$. $\hat{R}$ and $\hat{S}$ statistics below their critical values and $\hat{Q}$ statistics above its critical value indicate evidence of jumps.

Results The AJ test does not reject the null hypothesis at 5% significance level at any frequency considered, and this holds independently of the chosen starting point. For interval lengths up to 2 minutes, the $\hat{S}(4, 2, \delta)$ statistic is close to or above 2, the asymptotic limit under continuity. For 5 and 10 minute intervals, the statistic is for many starting points close to 1, the limit if there are jumps. For these longer time intervals, though, an even lower statistic is required for significance at the 5% level, since the resulting small number of increments reduces the power of the test. Large variability in test statistics over the starting point of the sample is seen in general, but the effect is most pronounced for longer time intervals, where some values are close to 1 and others far above 2.

Apart from a few significant values for 15 second intervals, the BNS test also does not show evidence against continuity for interval lengths up to 2 minutes. In general for these frequencies, the ratio statistic is close to 1, its asymptotic value under the null. For the two longest time intervals, the BNS test does show some evidence of jumps in the observed data series, as the statistic is significant for many starting points. For other starting points though, the statistic is close to 1, so the conclusion depends on where the sample is started.

The EV test statistic is significant at the 5% level for all frequencies and over all starting points, apart from a few values for 1 minute intervals. Accordingly, this test rejects the continuous model at the 5% level for practically all the specifications considered. Still, the values of the test statistic depend on the starting point. Notice that the pattern over starting points is similar to that for the BNS test for the two longest time intervals, recalling that respectively high and low values are critical for these two tests. For example, for 5 minute intervals, starting point t_0 in the range of $150 - 200$ seconds show least evidence of jumps, while values in the ranges $0 - 100$ and $250 - 300$ show more evidence of jumps for both tests.

5 Microstructure Noise

The results in the case study are difficult to reconcile with perfect discrete-time observations of a price that follows a process of the jump-diffusion type given in equation (1), regardless of whether the jump term is active or not. In particular is it difficult to explain the large variation in results over different sample starting points, and the fact that some statistics deviate from their asymptotic value under continuity in the opposite direction of what they should in the case of jumps. This section argues that noise in price observations may be the reason for these test results. To this end, a model is introduced, under which test statistics are likely to behave as observed in the case study. In the proposed model, noise arises from finite speed of market adjustment.

Underlying the model (1) is an assumption of perfect observations of prices as described by asset pricing theory. Observed transaction prices, though, are affected by frictions arising from the trading process, i.e. market microstructure effects. Examples of such effects include price discreteness and the bid-ask spread, but also effects arising from the manner in which information gets incorporated into prices, e.g., the way a market maker adjusts quotes taking the possibility of privately informed traders into account. Microstructure effects lead to transitory deviations from the efficient price, the price given by asset pricing theory. Therefore, in some applications the frictions may be disregarded, while in others, such as applications using high-frequency data, microstructure effects may be a first order important factor driving results.

A standard way to incorporate microstructure effects in observed prices is to separate them into a noise term around the efficient price described by asset pricing theory. In an additive noise model for log-prices, Y_t would thus be the sum of the efficient log-price, Y_t^*, and a noise term, ε_t, that captures the effect of microstructure,

$$Y_t = Y_t^* + \varepsilon_t. \tag{19}$$

Now, it is the efficient price, Y_t^*, that follows the model in (1), while different types of market frictions imply different statistical properties of the noise term, ε_t, regarding its own serial dependence and its relation to Y_t^*. This section first discusses microstructure effects related to the standard statistical assumption of white noise in observed prices. Then microstructure effects particularly relevant to the observation of discontinuous price changes is discussed, leading to a suggestion of related properties for the noise term in observed prices.

White Noise The simplest illustration of microstructure effects is white noise around the efficient price. Thus, ε_t in (19) has zero mean and is serially uncorrelated and independent of Y_t^*. Noise with these statistical properties may result from a setting with a market maker that maintains a bid-ask spread to cover inventory and order processing costs, as in the model by Roll (1984). When the spread is set only to cover operating costs, symmetrically around the efficient price, then random buy and sell trades at respectively the bid and ask prices will include noise that is uncorrelated over time and with the efficient price. Roll (1984) argued that these bounces between bid and ask prices would lead to negatively correlated observed returns and therefore to violation of the random walk model for observed prices.

White noise complicates inference about properties of Y_t^* from observed prices. The benefit of getting more frequent price observations no longer holds, since serially uncorrelated noise tends to become the dominant cause of price changes at high observation frequencies. That is, the size of the noise term in an observed δ-period price increment, $\varepsilon_t - \varepsilon_{t-\delta}$, doesn't decrease when the time interval is shortened, while $Y_t^* - Y_{t-\delta}^*$ goes to zero if there are no jumps in the interval. One can say that the signal relative to noise in observed price changes goes to zero as interval length goes to zero.

This effect is important in the related task of volatility estimation. Here, realized quadratic variation, a consistent estimator of integrated variance with perfect observations and no jumps, instead converges as $\delta \to 0$ to a term proportional to the variance of the noise term, if such is present, as shown by Zhang et al. (2005) and Bandi & Russell (2008). The straightforward solution is to choose interval lengths to balance noise effects against the advantages of frequent observations, e.g. by using 5 minute intervals as in Andersen, Bollerslev, Diebold & Labys (2001). Other methods that yield robust estimates using all data are argued to have better properties, though, such as the two-scale estimator of Zhang et al. (2005), and the pre-averaging approach of Jacod et al. (2007).

The possibility of microstructure effects leading to white noise in observed prices and its dominating effect on price changes at high observation frequencies is recognized in the jump testing literature. For example, Aït-Sahalia & Jacod (2009) show that their test has the limit $1/m$ when prices are observed with white noise. As this is even less than the no-noise asymptotic value with jumps, the test will reject continuity too often at high frequencies. The other tests considered here also rely on estimators of volatility that become biased at high frequencies when observations include white noise, and thus these test are likely to be affected, as well. As discussed in relation to volatility estimation, the issue can be dealt with by using lower observation frequency, so that noise is only a negligible part of observed price movements.

Alternatively, tests may be corrected to obtain robustness against i.i.d. noise, such as in Jiang & Oomen (2008), while Podolskij & Ziggel (2008) use the pre-averaging approach to construct a robust test.

The effect of a small white noise term on the jump tests considered is investigated further in the simulation study in the next section, along with the effect of prices being recorded discretely, i.e., in whole cents. The latter effect is similar to white noise in that it also becomes an increasing problem with higher observation frequency. To mitigate the effect of these noise types, tests in the GM case were discussed only for intervals longer than 15 seconds. The simulation study indicate that if shorter observation intervals are used then the actual size of tests starts to deviate from the chosen size when prices include a small white noise term or are observed in discrete values. Signs of these noise types were seen in the GM data. If tests were applied to intervals shorter than 15 seconds, all tests showed strong rejection of continuity, but this was not the case for intervals longer than 15 seconds for the AJ and BNS tests, as seen from results included in figure 3. Since statistics for 30 and 60 second intervals are practically never on the jump side of the value under continuity, the rejection at higher frequencies thus seems more likely to come from these noise types than just from higher power. These noise types, though, do not explain the unconventional values obtained for statistics at longer intervals, i.e., they do not explain why the AJ and to a lesser extent the BNS statistic often are larger than the their limit under continuity, whereas jumps actually should cause them to be lower. Further, nothing in these noise types has the potential to cause all statistics to vary over starting points to the extent observed.

Asymmetric Information Models More complicated microstructure effects than those leading to white noise in observed prices are necessary to explain results in the empirical case study. Fortunately in this regard, most microstructure literature argues that trades do not arrive independently of changes in the value of the asset. This allows for the possibility of microstructure effects related to changes in the efficient price. If changes in the underlying price affect the trade pattern, then the noise term in observed prices may not be white noise.

In asset pricing theory price changes are primarily driven by changes in information. Modelling the microstructure of how prices adjust to reflect new information through the trade process is therefore likely to relate microstructure effects to price changes. Such models necessarily must weaken to some extent the assumption of strong-form efficient markets that incorporate all information in prices. The first step is to let some agents have private information not reflected by market prices, while maintaining that prices reflect public information, i.e. are semi-strong form efficient. The models then study the process of how private information is learned by other agents through trades, leading to subsequent price adjustments. Classical among these asymmetric information models are the sequential trade models by Glosten & Milgrom (1985) and Easley & O'Hara (1992), which I discuss as an illustration of how price changes related microstructure effects may arise.

In the Roll (1984) model there is no private information, and the spread is set entirely to cover the market maker's operating costs. In the Glosten & Milgrom (1985) model on the other hand, some traders have information that is unknown by other liquidity traders and the market maker. As the marker maker doesn't know with which type of investor he trades, a spread is maintained to cover losses of trading with informed investors. That is, it covers adverse selection costs, as informed traders, in contrast to liquidity traders, only trade in the side of the market that is advantageous according to their superior information. There is information in trades then, and the likelihood for an uninformed trader that private information is positive or negative changes as trades arrive. The market maker takes this into account when setting bid and ask quotes by having these reflect current public information as well as the information revealed by another buy or sell. Gradually through the overweight of trades in one side of the market, the beliefs of the uninformed adjust toward private information, and quotes, and thus trade prices, converge to the valuation of the privately informed. Easley & O'Hara (1992) add event uncertainty to this setup by making it random whether an information event has occurred. As uninformed investors don't trade with some probability, and all investors are uninformed when no event has occurred, lack of trade reveals information by decreasing the likelihood that an event has happened. This model therefore can be used to connect trade volume to the amount of new information.

In sequential trade models where the spread is due to adverse selection costs alone, transaction prices are martingales with respect to public information. This follows as quotes set by the market maker include current public information plus information in respectively a buy or a sell trade. The trade price therefore changes as result of gradual increases in public information inferred from trades, or from externally arriving public information, which is assumed to immediately affect prices. Thus, the trade price process is a martingale with respect to public information. Indeed, the models only consider microstructure effects in relation to how private information gets incorporated into markets. Therefore their implications about deviations in market prices from the underlying private information price are particularly relevant. Arrival of new private information changes private valuation immediately, but only affects market prices gradually through the trading process as uninformed agents observe increased volume on one side of the market and quotes adjust accordingly. In relation to explaining empirical results of the jump tests, this illustrates effects that obstruct the otherwise discontinuous price change that would occur if all had received the information, and it illustrates a mechanism governing how market prices over time move toward a price otherwise reached instantaneously. To further consider mechanisms that limit the immediate market reaction and to also allow for effects in relation to new public information, I turn to models of gradual price adjustment.

Price adjustment This subsection discusses the adjustment of market prices to changes in the underlying efficient price. Inspiration is obtained from price adjustment models starting with Goldman & Beja (1979) and the microstructure effects discussed in this literature. In addition I discuss possible microstructure effects related to sudden changes in the underlying efficient price, e.g. arising due to new information.

In asset pricing theory prices are efficient with respect to publicly available information. Thus prices generate equilibrium between supply and demand of agents given available information and allocations that are efficient relative to agents' preferences and constraints. For prices that incorporate all available information changes must be unexpected, compensation for risk and time aside, and this leads to a martingale model for prices such as (1). When this model is used for observed prices, the assumption is that the mechanism through which market prices are formed does not significantly affect the properties of the model. As argued by Goldman & Beja (1979), for prices to be in the above described equilibrium at all times requires a strong price adjustment mechanism to changes in the environment, e.g. new information, that lead to changes in the efficient price. If this adjustment is less than perfect, discrepancies between market prices and efficient prices may exist temporarily. When jumps are thought of as changes in the environment, resulting in discontinuous changes in the efficient price, such discrepancies could affect jump tests. Though asset markets are very efficient, the question is whether the adjustment mechanism is so effective that discontinuous changes are perfectly observed at high-frequency.

Different elements of the adjustment mechanism, such as information dissemination, agent repositioning, and market clearing may all have effects which impede the immediate adjustment of market prices to changes in the efficient price. The speed with which information reaches agents may be limited, such that otherwise publicly available information relevant to the valuation of the asset is not immediately known to all agents. It may take time for agents to revalue assets in light of new information, for agents to chose their desired positions at different prices, and to update their buy and sell orders. These effects cause the order book to only gradually change in the direction of the full impact from released information. A specific consequence is that limit orders, which would have been removed by the originators had they been able to adjust immediately to new information, will instead generate trade. The time it takes for the market to move from one price level to another may be further affected by requirements on the market maker to maintain price continuity and avoid excessive price swings. In specialist markets, like the New York Stock Exchange, the market maker is usually required to maintain an orderly market, among other things meaning that the price should not be allowed to fluctuate dramatically from one trade to the next.[2]

The microstructure models discussed in the previous two subsections assume efficiency with

[2]In the NYSE rule 104: Dealings by Specialists, part 10, the specialist is required to ensure "the maintenaince of a fair and orderly market", after which it is stated that this "implies the maintenaince of price continuity".

respect to public information in the sense that quotes are immediately adjusted, and therefore trades only occur due to liquidity reasons or private information. Building on the asymmetric information models, Hasbrouck (1991) suggests a way to measure the information effect of trades. He argues that his method only measures private information if quotes indeed fully reflects public information and discusses several of the market clearing effects mentioned above as likely reasons that quote revision would be impaired.

The result of the discussed microstructure effects in relation to the price adjustment mechanism is to limit its speed to a gradual instead of immediate adjustment. Whenever discrepancies exist between the efficient price and the market price, agents will gradually update their orders, driving aggregate demand toward the perfect market level. Trades will be closed throughout this process due to stale orders and market maker smoothing, thus producing a gradual change in transaction prices. Goldman & Beja (1979) reasonably argue that the rate of price adjustment is increasing in the distance between market price and efficient price and propose that to a first approximation, this is a linear relation. This leads to the model

$$dY_t = \kappa \left(Y_t^* - Y_t\right) dt + \xi_t d\tilde{W}_t, \tag{20}$$

where Y_t is the quote midpoint and Y_t^* is the efficient price, both in logarithms. The diffusion term allows for some random imperfection in the adjustment process.

Different papers argue for the empirical relevance of a specification like (20). In an investigation of NYSE transaction data Hasbrouck & Ho (1987) find evidence of negative autocorrelation at first lag, but also positive autocorrelation at lags above one. They argue that a model specification similar to (20) together with a spread in the style of Roll (1984) can be used to generate the observed statistical properties. Amihud & Mendelson (1987) investigate the implications that different trading mechanisms have for price behavior. In dealership markets the market makers continuously post quotes at which they are willing to trade, while in clearing house markets limit and market orders are accumulated and cleared periodically. An important part of their comparison uses a model like (20) to discuss different statistical properties of returns in these two market types. Damodaran (1993) suggests a way to measure the price adjustment coefficient κ in (20) and finds evidence of lagged adjustment to new information in short period return intervals. Though not specifically related to an adjustment model, Hansen & Lunde (2006) more recently use a cointegration study of quotes and transaction prices from the TAQ data base for Dow Jones 30 stocks to recover the efficient price as the common stochastic trend. They find that noise in transaction prices is negatively related to efficient returns, which would be consistent with imperfect adjustment to changes in the efficient price.

Finite speed of market adjustment to news The model of Goldman & Beja (1979) focuses on general sluggishness in adjustment of market prices and less on the underlying types of changes in the efficient price. In fact, their formulation only includes continuous underlying changes. I instead wish to focus on microstructure effects that limit market adjustment speed when something has caused the efficient price to move discontinuously. The primary reason for discontinuous changes in the efficient price is arrival of new information relevant for valuation of the asset. Thus, the effects discussed in the previous subsection, stemming from limits to how quickly agents can obtain information and to how quickly they evaluate it and readjust their orders, are even more relevant here than in relation to continuous efficient price movements. The discussed impediments to market clearing are less related to the type of underlying movement, but nevertheless still relevant. In the change to the Goldman & Beja (1979) model that I propose, market prices adjust gradually to jump movements in the efficient price but instantaneously to continuous changes. This is justified by assuming that continuous changes don't arise due to externally arriving information that agents must analyze, but reflect small changes in the environment, such as different changes to individual agents that affect their demand. The remainder of the market is then assumed to be able to react without delay based on information revealed by price changes, without time-consuming analysis of external information being necessary.

The above discussion leads to a model for observed prices that adjust to new information with finite speed. The log of the underlying efficient price, Y_t^*, follows a jump-diffusion model as in (1), here written in stochastic differential equation form,

$$dY_t^* = \sigma_t dW_t + dJ_t^*. \tag{21}$$

The drift term has been excluded, since it will be negligible for the short time periods considered in the paper. An asterisk has been added to the compound Poisson jump term, J_t^*, to indicate that it's unobserved, similarly to the notation for Y. The behavior of the log of the market price[3], Y_t, is given by

$$dY_t = \sigma_t dW_t + \kappa\left(Y_t^* - Y_t\right)dt + \xi_t d\tilde{W}_t, \qquad Y_0 = Y_0^*. \tag{22}$$

The last two terms are similar to the model (20) of Goldman & Beja (1979). That is, the adjustment speed of Y_t toward Y_t^* is on average linear in the distance between these two prices, with the last diffusion term allowing for some imperfections in the adjustment process. κ is the coefficient determining the speed of market adjustment, while ξ_t, assumed to go to zero

[3]More precisely Y_t models the quote midpoint. Similarly to Hasbrouck & Ho (1987) a white noise term would have to be added to the adjustment model to obtain the bid-ask bounce effect in transaction prices described by Roll (1984). This could easily be added to the model, but the focus here is on jump related effects, while effects from the bid-ask bounce are mitigated by not using the highest data frequencies.

as Y_t tends to Y_t^*, governs how smooth the adjustment is. Relative to (20) it is assumed that continuous efficient price movements don't require adjustment, and thus $\sigma_t dW_t$ enters directly into the market price.

The adjustment model can be reformulated to emphasize that it's the jumps in efficient price that lead to gradual adjustment in market prices. By defining a term J_t as the observed part of jumps at time t, the model (21) - (22) can be written[4]

$$dY_t = \sigma_t dW_t + dJ_t, \tag{23}$$

$$dJ_t = \kappa \left(J_t^* - J_t\right) dt + \xi_t d\tilde{W}_t, \qquad J_0 = J_0^*. \tag{24}$$

This illustrates that when a jump occurs in J_t^*, it only becomes observed over time as the term J_t gradually moves toward J_t^* with randomness governed by ξ_t.

Alternatively, the model can be cast in the efficient price plus noise framework of (19),

$$\begin{aligned} Y_t &= Y_t^* + \varepsilon_t, \\ d\varepsilon_t &= -\kappa\varepsilon_t dt + \xi_t d\tilde{W}_t - dJ_t^* \qquad \varepsilon_0 = 0, \end{aligned} \tag{25}$$

where the efficient price still follows (21). Thus, noise that enters market prices shoots up in the opposite direction in the event of a jump in Y_t^*. It then reverts back toward zero at a speed proportional to the size of the noise itself with disturbances determined by ξ_t.

To illustrate the model, an example of how the noise term, ε_t, behaves is given in figure 5. A reasonable value for the adjustment speed parameter could be $\kappa = 50,000$, which implies that on average about 92% of a jump is observed after 5 minutes. ξ_t is assumed to be $\xi\varepsilon_t$ with $\xi = 50$. This gives significant disturbances in the adjustment initially, while also ensuring that these die out as the noise term, ε_t, goes to zero.

Jump test when market adjustment speed is finite Asymptotic identification of jumps relies on observing possible jumps, $Y_s^* - Y_{s-}^* \neq 0$, in single increments, y_i^δ, for $s \in (t_i - \delta, t_i]$, no matter how small $\delta > 0$ becomes. In the gradual price adjustment model this is impossible, and we cannot expect to get better power to detect underlying jumps by increasing the observation frequency. As clearly illustrated by (23) - (24), observed prices in this model follow a continuous process with a jump in the drift term. Properties of tests under continuity depend on the drift term being asymptotically negligible relative to the diffusion term. As the drift term, $\kappa\left(Y_t^* - Y_t\right)$, is very large following an underlying jump, the interval lengths that this holds for may be far shorter than empirically relevant intervals[5]. Thus, if observed prices are described

[4]From (21) and (23) $dY_t^* - dY_t = dJ_t^* - dJ_t$, and since $Y_0^* - Y_0 = J_0^* - J_0$, we have $Y_t^* - Y_t = J_t^* - J_t$.

[5]Consider 15 second intervals, parameter values used in the simulations, $\kappa = 50,000$, $\sqrt{\beta} = 0.4$, and $\xi = 50$, and a jump size at 2% of the price, $J_s^* - J_s = 0.02$. Then using an Euler discretization, the drift term will be about 8 bp. of the price, while a one standard deviation in respectively the $\sigma_t dW_t$ and $\xi\varepsilon_t d\hat{W}_t$ terms will be

by the gradual price adjustment model, tests may have problems detecting jumps, as they are not observed in single intervals, and they may behave quite differently from asymptotic properties under continuity, due to a temporarily very large drift term. Indeed, the latter leads to a series of consecutive discrete-time increments that are relatively large and in the same direction following an underlying jump. How this would affect the individual tests is now discussed.

A series of consecutive increments in the same direction increase $\hat{B}(p, 2\delta)_t$ relative to $\hat{B}(p, \delta)_t$, since $\left|y_i^{2\delta}\right|^p$ will then be large relative to $\left|y_i^{\delta}\right|^p + \left|y_{i-1}^{\delta}\right|^p$. For the AJ test this not only pushes $\hat{S}$ away from the jump value, 1, toward the continuous value, 2, but has the potential to drive it to even larger values.

The BNS test is based on the property that $\left|y_i^{\delta}\right| \left|y_{i-1}^{\delta}\right|$ will be small relative to $\left(y_i^{\delta}\right)^2$ when there is a jump in y_i^{δ}. If a jump is observed distributed across a series of consecutive changes in the same direction, then this effect will not be present, such that realized bipower variation, $\{Y_\delta\}_T^{[1,1]}$, will not be low relative to realized quadratic variation, $[Y_\delta]_T$. This may in finite samples even push the $\hat{R}$ statistic above 1, since $\left|y_i^{\delta}\right| \left|y_{i-1}^{\delta}\right|$ for the dominating increments around the jump will be of the same size as $\left(y_i^{\delta}\right)^2$, but the first is scaled by $\mu_1^{-2} > 1$ in the numerator of the ratio statistic, (7).

If a jump is spread across several increments due to noise, these will be smaller and whether the EV test then detects a jump depends on its magnitude and how it's split. Consecutive large increments increase the bipower variation estimate of integrated variance, but some of the now several jump increments may still stand out relative to estimated volatility over the entire local window.

The joint setting in section 1, under which the jump tests were introduced, followed Lee & Mykland (2008), while in Aït-Sahalia & Jacod (2009) and Barndorff-Nielsen & Shephard (2006) the drift term is only assumed to be cadlag. Thus, the jump term in the drift of the gradual price adjustment model doesn't satisfy the assumptions in section 1, but does satisfy those in the last two papers, in which it falls under the null hypothesis. A consequence of concluding that a continuous model is a good description of data is the ability to trade at all intermediate values when the price changes, which has implications e.g. for hedging and pricing. In the price adjustment model the reaction to an underlying jump is still quick, and the behavior after the jump is quite different from usual continuous movement, where the diffusion term dominates the drift at reasonably short intervals. Thus, if the purpose is to test whether a well behaved continuous model is sufficient to explain discrete-time observations, with its practical implications, it makes sense to exclude a model like the gradual adjustment model from the null hypothesis.

---

about 7 and 18 bp. of the price.

The GM case revisited A zoom on the crucial period of the GM case is given in the panels of figure 6, which show respectively price and accumulated volume from 12:00 to 12:15, the period with the large price decline. Instead of an immediate drop, the price changes gradually over about 5 minutes with many trades at intermediate price levels and with increasing volume. This behavior illustrates the proposed price adjustment model with a jump in underlying efficient price and gradual adjustment to the new level over a $5-8$ minute period. The observed increases in trade speed are also consistent with new information being released that leads investors to adjust their positions, and the sustained high volume indicates that this is a prolonged process.

Test results in the empirical case make sense when viewed in relation to how tests should behave in data from the price adjustment model, (21) and (22), with an active jump term. For intervals of 30 seconds to 2 minutes, the AJ statistic is larger than its asymptotic limit under continuity, 2, as it was argued to be if jumps are observed gradually. The BNS statistic is close to 1 and sometimes slightly above, which is also consistent with the proposed model. The EV test generally produces evidence against continuity in this interval range. As discussed above, against a price process that adjusts to underlying jumps in a gradual but quick fashion, the EV test may still have power to reject the null hypothesis of continuity.

When long intervals of 5 or 10 minutes are used, there is large variation in all three test statistics over the sample starting point, which matches gradual adjustment to an underlying jump that is cut differently into the longer increments. For 5 minute intervals, figure 7 illustrates the details of how the AJ and BNS statistics come out very differently for two different starting points. In one of the illustrated cases, the large price change arrives in a single increment, and thus $\left|y_i^\delta\right|^p + \left|y_{i+1}^\delta\right|^p$ has the same size as $\left|y_i^\delta + y_{i+1}^\delta\right|^p$, and $\left(y_i^\delta\right)^2$ is larger than $\mu_1^{-2}\left|y_i^\delta\right|\left|y_{i+1}^\delta\right|$, which implies that test statistics have values that indicate a jump. In the other case, the large price change is split in two increments, and thus $\left|y_i^\delta\right|^p + \left|y_{i+1}^\delta\right|^p$ is much smaller than $\left|y_i^\delta + y_{i+1}^\delta\right|^p$, while $\mu_1^{-2}\left|y_i^\delta\right|\left|y_{i+1}^\delta\right|$ is no longer smaller than $\left(y_i^\delta\right)^2$. This leads to a value indicating continuity and one that deviates from the continuity value in the opposite direction from the jump value. Thus, attempting to improve the ability to detect jumps by testing at longer intervals, where the full effect of a jump may be observed in a single interval, leads to dependence on the starting point.

Many of the effects that lead to gradual adjustment of the market price are probably more pronounced for surprising firm specific events than for regular macro announcements. Though news are released in either case, the market will be more ready to a fast reaction in the second case, while a longer and more gradual response is reasonable when not only the news, but also the release of news is a surprise. This would explain why the specific case of GM, where firm specific news were released at the day of the study, turns out to be a good illustration of the adjustment model, while one in other cases one may find market behavior with more readily observed discontinuous price movements.

6 Monte Carlo Study

The properties of the jump tests are now investigated in a Monte Carlo study. The first step is to consider performance under the no noise case, in which observed prices follow a jump-diffusion model satisfying (1). Both size, simulating under the null hypothesis, and power, simulating under the alternative, of the tests are checked. Next, I consider robustness toward white noise in the additive noise model (19) and price discreteness. Then the test performance in the price adjustment model (25) is investigated in more detail.

To mimic the form of the high-frequency data in the GM case, I simulate an intraday data series with observations every second. This gives 23,400 observations in each basic data set, which one can then subsample from in the same manner as in the empirical study. For each specification of the data generating process that I consider, a total of 10,000 sample paths are simulated. I report the mean and standard deviation of the test statistics over the simulations together with the observed relative rejection frequency of the null hypothesis for the tests at a significance level of 5%.

Continuous Price Behavior As the baseline model of continuous price behavior I use a stochastic volatility model in which the instantaneous variance follows a mean reverting square-root process,

$$\begin{aligned} dY_t &= \sigma_t dW_{1t}, \\ d\sigma_t^2 &= \omega\left(\beta - \sigma_t^2\right) dt + \gamma \sigma_t dW_{2t}. \end{aligned} \tag{26}$$

The drift has been set to zero, since this term will be negligible over the short time intervals to be studied. The Wiener processes are allowed to be correlated, satisfying $E\left[dW_{1t}dW_{2t}\right] = \rho dt$, which allows for the leverage effect, i.e. a negative relation between asset price and volatility. Since realizations for the process are generated as frequently as every second, I choose a simple Euler discretization to obtain the conditional distribution for $Y_t|Y_{t-\delta}$. Parameters are set to realistic values inspired by Aït-Sahalia & Jacod (2009), see notes to the tables for specific values.

For each simulation of the model in (26), the test is applied to increments obtained by sampling at different intervals lengths from 1 second to 10 minutes. The results in table 1 comply with the theoretical results for the test statistics under the null hypothesis. The $\hat{S}$ and $\hat{R}$ statistics are close to respectively 2 and 1, their asymptotic values under the null, and the average rejection rates of all three tests are close to the chosen 5% level. The AJ test rejects a little less than 5% of the time when the price is observed less frequently, while the other two tests reject slightly more often than the chosen size.

White Noise and Price Discreteness This section discusses the robustness of size results to respectively adding a small price independent noise term and to rounding off prices to whole cents. With respect to white noise an error term is added to prices generated according to (26) of size $0.5 \cdot 10^{-4}$, which is either positive or negative with equal probability. This illustrates the effect of a small bid-ask spread of 1 bp. of the price, i.e. 1 cent on a price level of \$100, and trades that are initiated by buyers or sellers with equal probability. With respect to price discreteness the prices generated by (26) are rounded off to one of the nearest 100'th decimals, randomly up or down with equal probability[6]. This thus also leads to a 1 cent spread with random trades at either side, but where it is included that prices must be quoted in whole cents. Test results for these two cases are shown in tables 2 and 3.

The white noise term leads to high rejection rates for the AJ test for 1 and 5 second intervals in line with the effect discussed theoretically by Aït-Sahalia & Jacod (2009). The BNS and the EV tests instead have lower size at the highest applied frequency, likely due to inflated estimates of volatility. For price discreteness both the AJ and BNS tests show too high rejection rates at 1 and 5 second intervals, while the EV test does not seem to be affected. The different effect on the BNS test likely arises as the spread that can only lie on whole cents leads to many increments with zero price changes, thus decreasing bipower variation relative to quadratic variation and indicating jumps. This did not happen in the pure white noise case, in which the underlying price moves the spread even if the next trades are on the same side of the marked.

The spread that the market maker maintains to cover operating costs may be larger than the 1 cent used in the simulations here and thus the effects may be stronger. Thus the results show that at least tests applied to intervals shorter than 15 seconds should be interpreted with care. These simulation results were the reason that the shortest interval length applied in the empirical case study was 15 seconds, though results for this frequency were still interpreted with the possibility of effects from these noise types in mind. In the study of the effects of other noise types in the simulations from now on white noise will not be added, but only intervals of at least 15 seconds will be considered.

Jumps in observed price The power to detect jumps is investigated by adding a compound Poisson term, J_t, to the baseline stochastic volatility model, such that the observed log-price follows

$$dY_t = \sigma_t dW_{1t} + dJ_t, \tag{27}$$

in which volatility is the same as in (26). Writing the jump component as $J_t = \sum_{j=1}^{N_t} c_j$, the Poisson process N_t has intensity λ, and the distribution of jump sizes, c_j, is uniform on

[6]Generated prices are in logs, but the round off is done in levels and then transfered back to logs to get data for the tests.

$c\,([-2,-1] \cup [1,2])$. Here, c is a constant chosen to set the total variance from jumps, $\lambda\,(7/3)\,c^2$, at some fraction, η, of average variance from the continuous part, which is approximately equal to β.

The jump size distribution ensures that jumps are bounded away from zero, and the determination of c lets frequent jumps be small relative to infrequent jumps. With variance of the jump term at half of average continuous variance, $\eta = 0.5$, jumps at the two considered intensities, $\lambda = 1$ and $\lambda = 10$, are of absolute sizes in respectively the range $[1.17, 2.33]\,\%$ and $[.37, .74]\,\%$ of the price. Only simulated paths with $N_t > 0$ are used, as we cannot distinguish the case with $N_t = 0$ from continuity, see the discussion below (4). This implies that λ is slightly higher in the resulting set of samples paths. Results of the simulations are in table 4.

For series with few jumps, $\lambda = 1$, all tests have power close to 1 for short time intervals. The power of tests decreases with the length of the intervals, but this happens faster for the AJ test compared to the other two, and slightly faster for the BNS test relative to the EV test. When smaller but more frequent jumps are considered, the power of the AJ test disappears for intervals longer than 15 seconds, while the BNS and EV tests retain high power for interval lengths up to about 2 minutes.

Gradual price adjustment Data is simulated from the gradual price adjustment model to check the power of the tests against this alternative to the standard continuous model. The additive noise term in observed prices thus follows (25), such that jumps in the observed series are realized gradually as in (23)-(24). The speed of market adjustment parameter κ is set to 50,000, implying that about 92% of the jump impact is observed after 5 minutes, and the adjustment disturbance parameter ξ is set to 50. An example of a simulated path for observed prices is shown in figure 8, and results for jump tests at different frequencies are in table 5.

Results confirm that the tests have difficulties rejecting the null hypothesis when the deviation is due to jumps in the underlying price that are not fully observed in a single price increment. First, note that the $\hat{S}$ statistic is pushed above 2, the asymptotic value under the null, instead of toward 1, the asymptotic value with jumps. This effect was explained in the previous section and is caused by large neighboring increments in the same direction. For similar reasons, the $\hat{R}$ statistic is close to and sometimes above 1 for shorter intervals. The EV statistic should be increasing for shorter intervals but has largest mean for δ at 2 minutes, since noise then splits the jump into several observed increments, decreasing the size of the largest one. To sum up, the AJ test practically never rejects the null, the BNS test does so about 40% of the time for long intervals, 5 to 10 minutes, where most of the jump is allowed to impact observed prices, while the EV test rejects continuity for a large fraction of simulated paths all the way down to 15 second intervals. Results thus show that the EV test has most power of the three tests considered against this alternative to the standard continuous model.

Robustness to Starting Point In the empirical case study test results for longer time intervals were very dependent on the starting point, t_0. This is now investigated in the simulation study by repeating for different starting values the tests in the two cases with jumps, respectively with and without limits to price adjustment. For the results in table 6 and 7, the reported means and standard deviations are respectively the mean over both starting values and simulations, and the square root of the average variance over starting values. The rejection rates give the frequencies of simulated series for which at least 95%, respectively 50%, of the starting values lead to rejection of the null at 5% significance level. These rejection rates can help judge whether rejection rates in table 4 and 5 for constant t_0 result mainly from variations over different realizations of the randomness, or mainly from variation in the starting point relative to the jump time.

Results for the reference case without noise in observations are in table 6. Just requiring rejection for most starting points, more than 50% of t_0's, gives rates similar to those in table 4, which is expected, since $t_0 = 0$ used there is also random relative to the jump time. Requiring rejection almost independently of the starting point, for at least 95% of the t_0's, reduces rejection rates considerably in cases where the power in table 4 was not close to one. This shows that the rejection rates for longer time intervals in table 4 are not only a result of variation of different realizations of randomness, but also depends on how observations fall relative to the jump, even when the jump is observed in a single increment.

Results for the case with gradual price adjustment are in table 7. Clearly, tests that almost never reject continuity for single starting points in table 5 will also not be able to do this when also some uniformity over starting points is required. Rejection rates for the BNS test for intervals of length 5 to 10 minutes fall from about 40% almost to 0% when respectively 50% and 95% uniformity over starting points is required. This shows that the indication in table 5 that the BNS test despite noise still has reasonable a 40% ability to detect jumps if long enough intervals are used did not arise as a result of rejecting continuity almost independently of the starting point for this fraction of sample paths, but from starting at the right point relative to the jump. The rejection rates for the EV test also decline when stability over t_0 is required, especially so for the longer intervals. At medium length intervals, though, the test still has power to detect jumps in about 50% of the simulated series almost independently of the starting point.

7 Conclusion

This paper demonstrates the difficulty in detecting jumps if microstructure effects cause discontinuous changes in the underlying price to affect observed prices only in a gradual fashion over time, and a model of prices when such effects are present is suggested. A case study illustrates how a negative news announcement is followed in the market by a period with many small price changes, almost only in the same direction, instead of by an immediate adjustment between two observed prices. Though gradual price adjustment makes observed prices move continuously, the price behavior following underlying jumps with many moves almost uniquely in the same direction are very unlikely in standard diffusion models and thus has different implications for practical purposes related to the price behavior. Therefore, although the alternative is not a perfectly observed jump, we would still like tests to tell us that the standard continuous diffusion model is not an adequate description of data.

As jump tests depend on detecting the jump in a single increment, it is no surprise that observing jumps gradually makes it difficult for tests to detect the deviation from usual continuous behavior. There are differences among tests, though, as tests that build their inference on comparisons of neighboring increments have the largest difficulties. The extreme value test based on Lee & Mykland (2008) relies less on such comparisons and thus has better power to detect this type of deviation from normal continuous price behavior.

If jumps are not observed immediately at high frequency, a simple suggestion would be to base tests on observations at a frequency for which jumps have time to fully impact observed prices. This solution is shown to be flawed, as the low frequency observations can be extracted from the full record of observations in many different ways. Test results will then be very dependent on the starting point from which increments are calculated, as shown both in the empirical case and in the simulation study.

It would be relevant to further study the generality of whether reactions to news are observed gradually in high frequency data due to microstructure effects. Differences may be expected over markets and types of news announcements, e.g. this pattern may be more common for specific news concerning single stocks relative to news released in regular macro announcements, where the release of news is expected. It would also be relevant to develop a way to remove the randomness in outcomes arising due to many possible points in the full record of observations from which increments at lower frequency can be calculated. A method could be to repeat the test for all possible starting points at the chosen frequency and then use a proper aggregation method to arrive at a single test for this frequency. Alternatively, one could test at high frequency with adjustments of the tests that account for noise due to microstructure effects.

Appendix

Extreme Value Theory

Consider a sample of i.i.d. random variables $\{X_1, ..., X_n\}$ with symmetric distribution around zero described by the distribution function F. The statistic that counts how many X variables are numerically larger than a threshold u is

$$H_u(n) = \sum_{i=1}^{n} I\{|X_i| > u\}.$$

Due to the symmetric distribution, $P(|X_i| > u) = 2(1 - F(u))$, and since the X_i's are independent $H_u(n) \sim Bin(n, 2(1 - F(u)))$. For large n and small p the Binomial distribution is very well approximated by the Poisson distribution, so it may sometimes be convenient to use $H_u(n) \overset{a}{\sim} Poi(2n(1 - F(u)))$.

Name the k'th largest $|X_i|$ variable $Q_k(n)$. The relation between this extreme order statistic and the number of exceedences over a threshold is described by equivalence of the events {the k'th largest value less than u} and {at most $k-1$ values larger than u}, i.e.,

$$\{Q_k(n) \leq u\} = \{H_u(n) \leq k - 1\}. \tag{28}$$

From this it follows immediately that the probability that the k'th largest absolute value is less than u can be written as the binomial probability of up to $k-1$ successes in n trials with probability $p = 2(1 - F(u))$,

$$P(Q_k(n) \leq u) = P(H_u(n) \leq k - 1) = \sum_{r=0}^{k-1} \binom{n}{r} 2^r [1 - F(u)]^r [2F(u) - 1]^{n-r}.$$

Define the value that the k'th largest variable doesn't exceed with probability $1 - \alpha$ as

$$C_Q^{\alpha}(k) = \{u : P(Q_k(n) \leq u) = 1 - \alpha\}. \tag{29}$$

Since the extreme value test statistic $\hat{Q}(\delta)$ in (18) approximately follows the same distribution as $Q_1(n)$ under the null hypothesis, the value (29) will be the critical value at α significance level for tests against alternatives where high extremes are critical to the null hypothesis.

Jump Term Parameters

J_t is a compound Poisson process, $J_t = \sum_{j=1}^{N_t} c_j$, where the Poisson process N_t has intensity λ at 1 or 10 per day, and jump sizes c_j are uniform on $c([-2, -1] \cup [1, 2])$. Here, c is a constant chosen to set the total variance from jumps, $\lambda(7/3)c^2$, at a fraction, $\eta = 0.5$, of average variance from the continuous part, which is approximately $\beta = 0.16$. The way c is chosen implies that jumps tend to be smaller when they are more frequent.

For c_j uniform on $c\left([-\theta,-1]\cup[1,\theta]\right)$, the variance of jump sizes is

$$var\left(c_j\right) = 2\int_c^{\theta c} \frac{v^2}{2c\left(\theta-1\right)}dv = \frac{1}{c\left(\theta-1\right)}\left[\frac{v^3}{3}\right]_c^{\theta c} = \frac{c^2}{3}\frac{\theta^3-1}{\theta-1} = c^2\left(\theta^2+\theta+1\right)/3.$$

This implies that the jump variance in a small increment is

$$var\left(J_\delta\right) = \delta\lambda c^2\left(\theta^2+\theta+1\right)/3,$$

and specifically for $\theta = 2$ the jump variance per time unit is $c^2\lambda\left(7/3\right)$. The average variance of the continuous part per time unit is approximately the level of mean reversion for σ^2 which is β. Therefore, to set jump variance at η fraction of continuous variance, c must be set to

$$\begin{aligned} \eta\beta &= \lambda c^2\left(\theta^2+\theta+1\right)/3 \\ c &= \sqrt{3\eta\beta/\left[\lambda\left(\theta^2+\theta+1\right)\right]}. \end{aligned}$$

References

Aït-Sahalia, Y. & Jacod, J. (2009), 'Testing for jumps in a discretely observed process', *Forthcoming in Annals of Statistics* .

Amihud, Y. & Mendelson, H. (1987), 'Trading mechanisms and stock returns: An empirical investigation', *The Journal of Finance* **42**(3), 533–553.

Andersen, T. G., Bollerslev, T., Diebold, F. X. & Ebens, H. (2001), 'The distribution of realized stock return volatility', *Journal of Financial Economics* **61**, 43–76.

Andersen, T. G., Bollerslev, T., Diebold, F. X. & Labys, P. (2001), 'The distribution of realized exchange rate volatility', *Journal of the American Statistical Association* **96**(453), 42–55.

Bakshi, G., Cao, C. & Chen, Z. (1997), 'Empirical performance of alternative option pricing models', *The Journal of Finance* **52**(5), 2003–2049.

Bandi, F. & Russell, J. (2008), 'Microstructure noise, realized variance, and optimal sampling', *The Review of Economic Studies* **75**, 339–369.

Barndorff-Nielsen, O. & Shephard, N. (2004), 'Power and bipower variation with stochastic volatility and jumps', *Journal of Financial Econometrics* **2**(1), 1–37.

Barndorff-Nielsen, O. & Shephard, N. (2006), 'Econometrics of testing for jumps in financial economics using bipower variation', *Journal of Financial Econometrics* **4**(1), 1–30.

Brownlees, C. T. & Gallo, G. (2006), 'Financial econometric analysis at ultra-high frequency: Data handling concerns', *Computational Statistics and Data Analysis* **51**, 2232–2245.

Damodaran, A. (1993), 'A simple measure of price adjustment coefficients', *The Journal of Finance* **48**(1), 387–400.

Easley, D. & O'Hara, M. (1992), 'Time ans the process af security price adjustment', *The Journal of Finance* **47**(2), 577–605.

Glosten, L. R. & Milgrom, P. R. (1985), 'Bid, ask and transaction prices in a specialist marker with heterogeneously informed traders', *Journal of Financial Economics* **14**, 71–100.

Goldman, M. B. & Beja, A. (1979), 'Market prices vs. equilibrium prices: Returns' variance, serial correlation, and the role of the specialist', *The Journal of Finance* **34**(3), 595–607.

Hansen, P. R. & Lunde, A. (2006), 'Realized variance and market microstructure noise', *Journal of Business and Economic Statistics* **24**(2), 127–161.

Hasbrouck, J. (1991), 'Measuring the information content of stock trades', *The Journal of Finance* **46**(1), 179–207.

Hasbrouck, J. & Ho, T. S. Y. (1987), 'Order arrival, qoute behavior, and the return-generating process', *The Journal of Finance* **42**(4), 1035–1048.

Jacod, J., Li, Y., Mykland, P. A., Podolskij, M. & Vetter, M. (2007), 'Microstructure noise in the continuous case: The pre-averaging approach', *Working Paper* .

Jiang, G. & Oomen, R. (2008), 'Testing for jumps when asset prices are observed with noise - a "swap variance" approach', *Journal of Econometrics* **144**, 352–370.

Johannes, M. (2004), 'The statistical and economic role of jumps in continuous-time interest rate models', *The Journal of Finance* **59**(1), 227–260.

Lee, S. & Mykland, P. (2008), 'Jumps in financial markets: A new nonparametric test and jump dynamics', *Review of Financial Studies* **21**(6), 2535–2563.

Mancini, C. (2004), 'Estimation of the characteristics of the jumps of a general poisson-diffusion model', *Scandinavian Actuarial Journal* **2004**(1), 42–52.

Mancini, C. (2006), 'Non paramteric threshold estimation for models with stochastic diffusion coefficient and jumps', *Working Paper* .

Mancini, C. & Reno, R. (2006), 'Threshold estimation of jump-diffusion models and interest rate modelling', *Working Paper* .

Merton, R. (1976), 'Option pricing when underlying stock returns are discontinuous', *Journal of Financial Economics* **3**, 125–144.

Naik, V. & Lee, M. (1990), 'General equilibrium pricing of options on the market portfolio with discontinuous returns', *The Review of Financial Studies* **3**(4), 493–521.

Podolskij, M. & Ziggel, D. (2008), 'New tests for jumps: A threshold-based approach', *Working Paper* .

Roll, R. (1984), 'A simple implicit measure of the effective bid-ask spread in an efficient market', *The Journal of Finance* **39**(4), 1127–1139.

Zhang, L., Mykland, P. A. & Aït-Sahalia, Y. (2005), 'A tale of two time scales: Determining integrated volatility with noisy high frequency data', *American Statistical Association* **100**(472), 1394–1411.

Table 1:

Levels of tests. Simulation under the null hypothesis.

		Mean and std. deviation			*Rejection rate,* $\alpha = 5\%$		
δ	n	$\hat{S}$	$\hat{R}$	$\hat{Q}$	*AJ*	*BNS*	*EV*
1 sec	23,400	2.000 .048	1.000 .005	4.219 .285	.049	.049	.052
5 sec	4,680	1.999 .107	1.000 .012	3.845 .311	.043	.057	.055
15 sec	1,560	2.001 .188	.999 .020	3.561 .335	.046	.057	.055
30 sec	780	2.002 .262	.999 .028	3.371 .347	.038	.055	.054
1 min	390	1.996 .361	.997 .040	3.172 .367	.039	.058	.055
2 min	156	2.004 .565	.993 .062	2.896 .395	.033	.064	.055
5 min	78	2.035 .783	.988 .088	2.670 .412	.029	.068	.053
10 min	39	1.966 1.023	.973 .122	2.432 .440	.032	.079	.059

Note: This table reports results from 10,000 simulations of the AJ, BNS, and EV tests under the null hypothesis of no jumps. The means and standard deviations of the test statistics over simulations are shown together with the relative frequencies at which the test statistics are critical to the null at 5% significance level. The data generating process is the stochastic volatility model: $dY_t = \sigma_t dW_{1t}$, in which squared volatility follows $d\sigma_t^2 = \omega\left(\beta - \sigma_t^2\right) dt + \gamma\sigma_t dW_{2t}$, and $E\left[dW_{1t}dW_{2t}\right] = \rho dt$. Parameters are set to realistic values for stocks, $\sqrt{\beta} = 0.4$, $\gamma = 0.5$, $\omega = 5$, and $\rho = -0.5$, following Aït-Sahalia & Jacod (2009). In calculation of realized truncated p'th variation, necessary for critical values in the AJ test, parameters are set to $\psi = 4\sqrt{\beta}$ and $\varpi = 0.47$.

Table 2:

Levels of tests when prices include white noise.

		Mean and std. deviation			*Rejection rate,* $\alpha = 5\%$		
δ	n	$\hat{S}$	$\hat{R}$	$\hat{Q}$	AJ	BNS	EV
1 sec	23,400	1.711	1.002	4.164	1.000	.021	.035
		.044	.005	.272			
5 sec	4,680	1.931	1.000	3.838	.158	.051	.052
		.105	.012	.310			
15 sec	1,560	1.975	.999	3.559	.054	.057	.055
		.181	.020	.333			
30 sec	780	1.991	.999	3.368	.045	.056	.052
		.260	.028	.350			
1 min	390	1.998	.998	3.168	.039	.057	.054
		.363	.040	.366			

Note: This table reports results from 10,000 simulations of the AJ, BNS, and EV tests under the null hypothesis of no jumps, but when prices include white noise. The data generating process for the raw prices is the same as that used to generate results in table 1, but an error term of $.5 \cdot 10^{-4}$ is added or subtracted randomly with equal probability for each observation. This is meant to reflect a small bid-ask spread of 1 bp. of the price, i.e. 1 cent on a price of \$100, and trades that are randomly initiated by either a buyer or seller.

Table 3:

Levels of tests when prices are observed in discrete values.

		Mean and std. deviation			*Rejection rate,* $\alpha = 5\%$		
δ	n	$\hat{S}$	$\hat{R}$	$\hat{Q}$	*AJ*	*BNS*	*EV*
1 sec	23,400	1.634	.955	4.255	1.000	1.000	.052
		.043	.006	.292			
5 sec	4,680	1.908	.988	3.865	.212	.275	.061
		.103	.012	.317			
15 sec	1,560	1.965	.995	3.575	.059	.084	.059
		.181	.020	.336			
30 sec	780	1.988	.997	3.373	.047	.065	.053
		.260	.028	.346			
1 min	390	1.996	.996	3.174	.035	.063	.056
		.361	.039	.366			

Note: This table reports results from 10,000 simulations of the AJ, BNS, and EV tests under the null hypothesis of no jumps, but when prices are observed in discrete values. The data generating process for the prices is the same at that used to generate results in table 1, but prices are rounded off whole cents, randomly up or down with equal probability, to reflect a trade initiated by either a buyer or seller.

Table 4:

Ability to reject the null hypothesis when paths contain jumps.

		Mean and std. deviation			*Rejection rate,* $\alpha = 5\%$		
δ	n	$\hat{S}$	$\hat{R}$	$\hat{Q}$	*AJ*	*BNS*	*EV*
				$\lambda = 1$ *jump per day*			
15 sec	1,560	1.018 .147	.626 .126	26.293 4.721	.999	1.000	1.000
30 sec	780	1.036 .220	.639 .126	18.589 3.427	.987	1.000	1.000
1 min	390	1.067 .293	.659 .126	13.104 2.509	.894	.995	1.000
2 min	156	1.129 .430	.686 .130	9.244 1.878	.624	.953	.999
5 min	78	1.300 .682	.730 .145	5.860 1.371	.163	.784	.967
				$\lambda = 10$ *jumps per day*			
15 sec	1,560	1.143 .210	.759 .064	10.800 1.202	.928	.999	1.000
30 sec	780	1.257 .308	.789 .064	7.805 1.028	.274	.994	1.000
1 min	390	1.414 .410	.828 .066	5.741 .892	.070	.945	.997
2 min	156	1.598 .546	.872 .076	4.324 .788	.035	.661	.827
5 min	78	1.828 .785	.920 .101	3.183 .633	.022	.247	.301

Note: This table reports results from 10,000 simulations of the AJ, BNS, and EV tests when the paths include jumps. The data generating process is $dY_t = \sigma_t dW_{1t} + dJ_t$, where σ_t follows the same stochastic volatility process used in table 1. J_t is a compound Poisson process, $J_t = \sum_{j=1}^{N_t} c_j$, where the Poisson process N_t has intensity λ at 1 or 10 per day, and jump sizes c_j are uniform on $c\,([-2,-1] \cup [1,2])$. Here, c is a constant chosen to set the total variance from jumps, $\lambda\,(7/3)\,c^2$, at some fraction, $\eta = 0.5$, of average variance from the continuous part, which is approximately $\beta = 0.16$. The way c is chosen implies that jumps tend to be smaller when they are more frequent. Simulated paths without jumps are removed as tests don't have power against potential but unrealized jumps.

Table 5:

Rejection rates in gradual price adjustment model

		Mean and std. deviation			*Rejection rate,* $\alpha = 5\%$		
δ	n	$\hat{S}$	$\hat{R}$	$\hat{Q}$	AJ	BNS	EV
15 sec	1560	2.555 .611	1.004 .022	4.866 1.094	.012	.041	.704
30 sec	780	2.864 .876	1.007 .033	4.847 1.146	.013	.047	.747
1 min	390	3.076 1.149	1.003 .048	4.860 1.184	.021	.065	.800
2 min	156	2.785 1.304	.968 .075	4.880 1.243	.039	.184	.837
5 min	78	2.132 1.228	.875 .133	4.474 1.222	.040	.404	.793
10 min	39	1.783 1.166	.834 .166	3.659 1.032	.028	.393	.635

Note: This table reports results from 10,000 simulations of the AJ, BNS, and EV tests when the price follows the gradual adjustment model. Thus, observed log-prices are $Y_t = Y_t^* + \varepsilon_t$, where Y_t^* follows the data generating process in table 4 with jump intensity $\lambda = 1$ per day, and the noise term satisfies (25), $d\varepsilon_t = -\kappa\varepsilon_t dt + \xi\varepsilon_t d\tilde{W}_t - dJ_t^*$. Parameters are set to $\kappa = 50,000$, corresponding to approximately 8% noise left after 5 minutes, and $\xi = 50$, introducing some randomness in the way jumps disseminate to observed prices.

Table 6:

Ability to reject the null hypothesis when paths contain jumps. Tests are repeated for different sample starting points.

		Mean and std. deviation			*Rejection for min.* 95% (50%) *of* t_0*'s*, $\alpha = 5\%$		
δ	n	$\hat{S}$	$\hat{R}$	$\hat{Q}$	AJ	BNS	EV
15 sec	1,560	1.018 .126	.626 .020	26.220 1.344	.995 1.000	.999 1.000	.999 1.000
30 sec	780	1.038 .187	.640 .028	18.513 1.075	.915 .999	.997 .999	.999 .999
1 min	390	1.074 .260	.660 .038	13.062 .913	.525 .966	.967 .997	.997 .998
2 min	156	1.137 .363	.686 .053	9.210 .787	.088 .680	.845 .969	.992 .997
5 min	78	1.297 .571	.732 .081	5.802 .671	.001 .071	.473 .807	.892 .970
10 min	39	1.474 .787	.774 .111	4.032 .619	.000 .001	.146 .539	.495 .792

Note: This table reports results from 10,000 simulations of the AJ, BNS, and EV tests when paths include jumps. The data generating process is the same as that used in table 4 with jump intensity λ at 1 per day. The sensitivity of the results to changing the point from which increments are calculated is illustrated. Thus, for each simulation of a price path, the tests using increments at a given time interval, δ, are performed for all starting values that lead to a different set of increments. The mean is taken over both simulations and starting points, while the standard deviation is the square root of the simulation mean of variances over starting points. Rejection rates indicate the number of simulations for which the null hypothesis is rejected at the 5% level for at least 95%, respectively 50%, of the starting values.

Table 7:

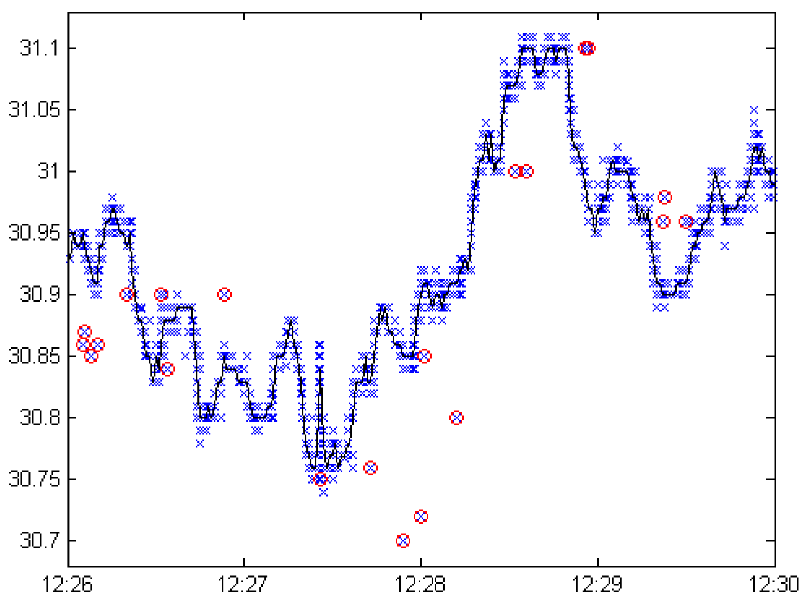
Rejection rates in gradual price adjustment model
Tests are repeated for different sample starting points.

		Mean and std. deviation			*Rejection for min.* 95% (50%) *of* t_0*'s*, $\alpha = 5\%$		
δ	n	$\hat{S}$	$\hat{R}$	$\hat{Q}$	AJ	BNS	EV
15 sec	1560	2.551 .438	1.003 .019	4.875 .672	.000 .000	.000 .003	.331 .741
30 sec	780	2.878 .672	1.007 .026	4.848 .677	.000 .000	.000 .004	.423 .782
1 min	390	3.052 .972	1.003 .038	4.877 .718	.000 .000	.000 .008	.495 .841
2 min	156	2.794 1.178	.968 .060	4.872 .814	.000 .001	.002 .083	.527 .885
5 min	78	2.099 1.111	.876 .107	4.453 .873	.000 .001	.018 .372	.334 .869
10 min	39	1.834 1.040	.836 .130	3.613 .720	.000 .000	.013 .363	.084 .697

Note: This table reports results from 10,000 simulations of the AJ, BNS, and EV tests when prices follow the gradual adjustment model. The specifications are the same as in table 5, but the sensitivity of results to changing the point from which increments are calculated is illustrated. Thus, for each simulation of a price path, the tests using increments at a given time interval, δ, are performed for all starting values that lead to a different set of increments. The mean is taken over both simulations and starting points, while the standard deviation is the square root of the simulation mean of variances over starting points. Rejection rates indicate the number of simulations for which the null hypothesis is rejected at the 5% level for at least 95%, respectively 50%, of the starting values.

Figure 1:

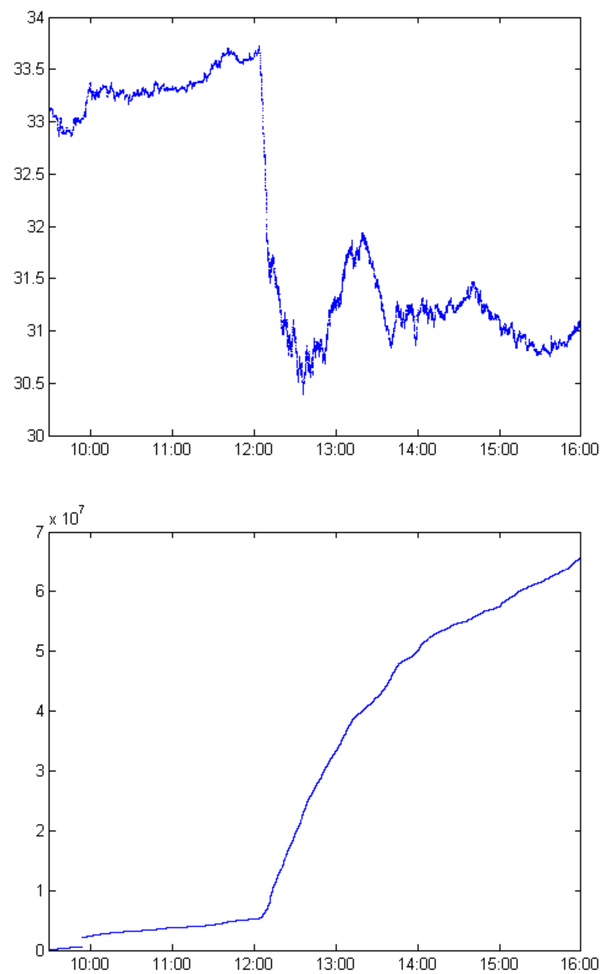
Illustration of the Filtering Procedure



Note: The crosses mark all recorded transaction prices, while those with a circle around are removed by the filter. The dark line is the median of remaining transaction prices for each second. This is the series used for the jump tests.

Figure 2:

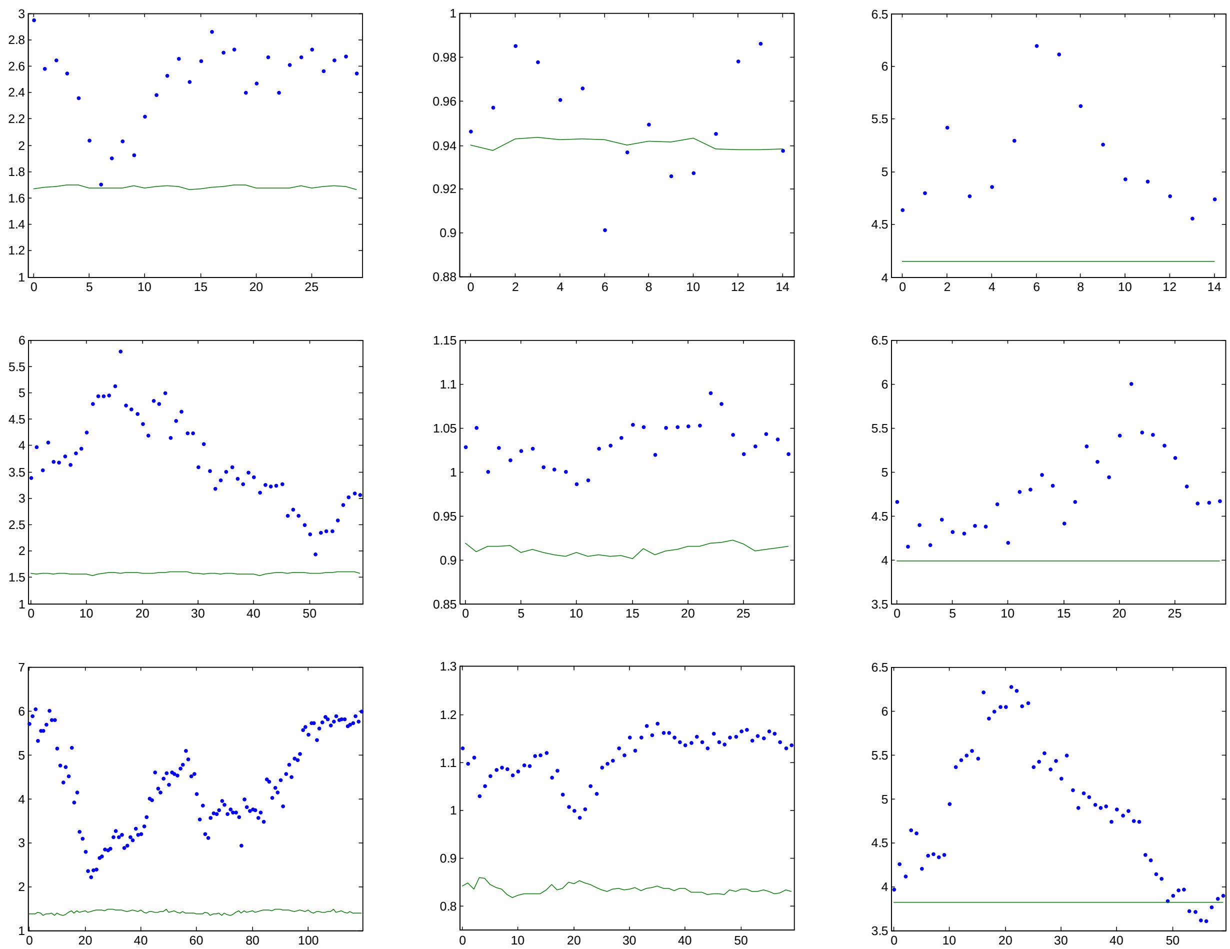
General Motors, October 6, 2006.



Note: The top figure shows observed transaction prices, while the bottom figure shows accumulated volume over the trading day.

Figure 3:

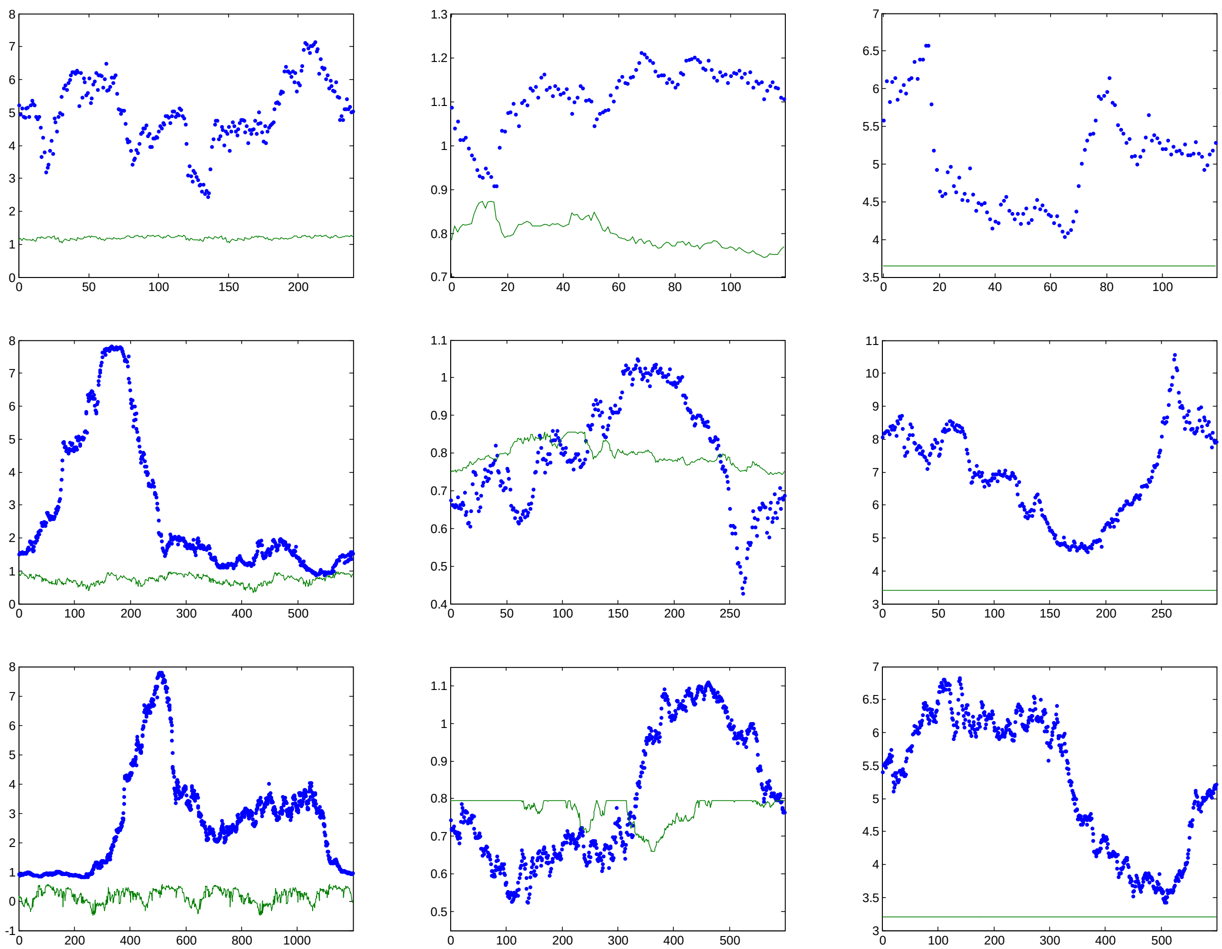
Jump tests in GM data sampled at 15 *and* 30 *second as well as* 1 *minute intervals.*



Note: Each panel shows results of a given test applied to increments at a given interval length. The rows are different time intervals: 15 and 30 seconds, and then 1 minute intervals, from top to bottom. The columns are the AJ, BNS and EV tests, from left to right. For each test at a given frequency results are shown for all relevant different sample starting values. The dots are values of the test statistics, $\hat{S}(2,4,\delta)$, $\hat{R}(\delta)$ and $\hat{Q}(\delta)$ respectively, and the lines are critical values at 5% significance level. The AJ and BNS tests must be less than their critical values to be critical for the null, while the EV test must exceed it.

Figure 4:

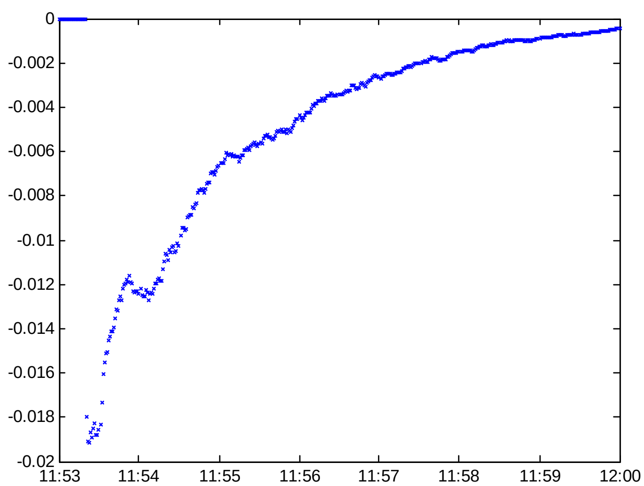
Jump tests in GM data sampled at 2, 5, and 10 minute intervals.



Note: Each panel shows results of a given test applied to increments at a given interval length. The rows are different time intervals: 2, 5, and 10 minute intervals, from top to bottom. The columns are the AJ, BNS and EV tests, from left to right. For each test at a given frequency results are shown for all relevant different sample starting values. The dots are values of the test statistics, $\hat{S}(2, 4, \delta)$, $\hat{R}(\delta)$ and $\hat{Q}(\delta)$ respectively, and the lines are critical values at 5% significance level. The AJ and BNS tests must be less than their critical values to be critical for the null, while the EV test must exceed it.

Figure 5:

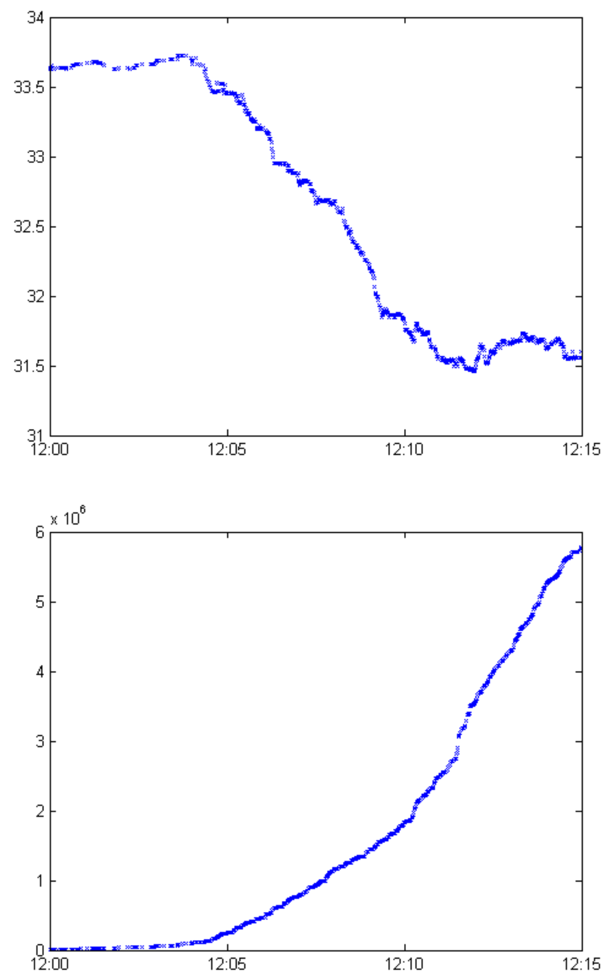
Illustration of Jump Noise



Note: Example of a path for the noise term in the gradual price adjustment model. The graph shows how much the observed price, Y_t, deviates from the underlying price Y_t^*, following a jump in Y_t^*. The noise term is given by $d\varepsilon_t = -\kappa\varepsilon_t dt + \xi\varepsilon_t d\tilde{W}_t - dJ_t^*$, where parameters are set to $\kappa = 50,000$ and $\xi = 50$. As the model is for log-prices the jump corresponds to about 2% of the price level.

Figure 6:

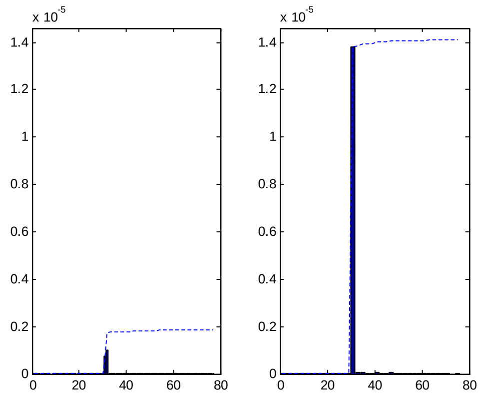
General Motors, $12:00$ *to* $12:15$*, October* $6th$*,* 2006*.*



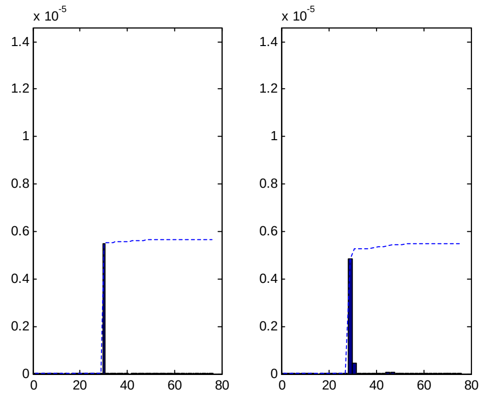
Note: Top figure shows observed transaction prices, while the bottom figure shows accumulated volume during the selected period.

Figure 7:

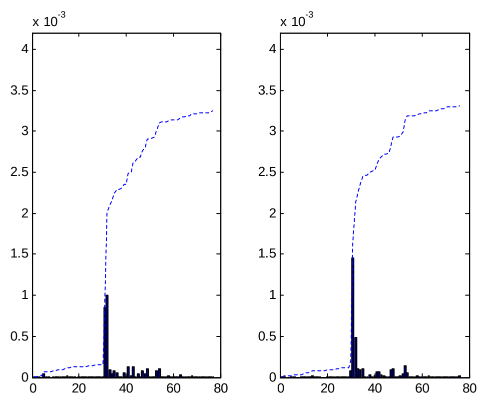
Illustration of variation in test statistics over starting points



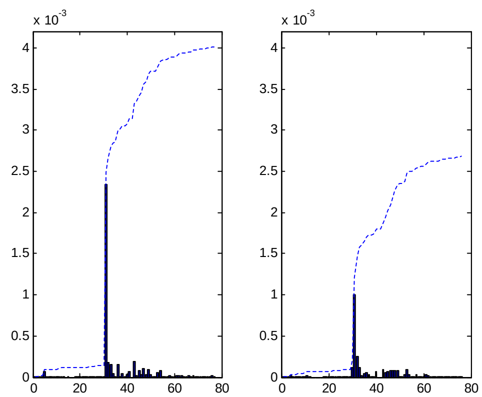
AJ, $t_0 = 170s$



AJ, $t_0 = 550s$



BNS, $t_0 = 170s$

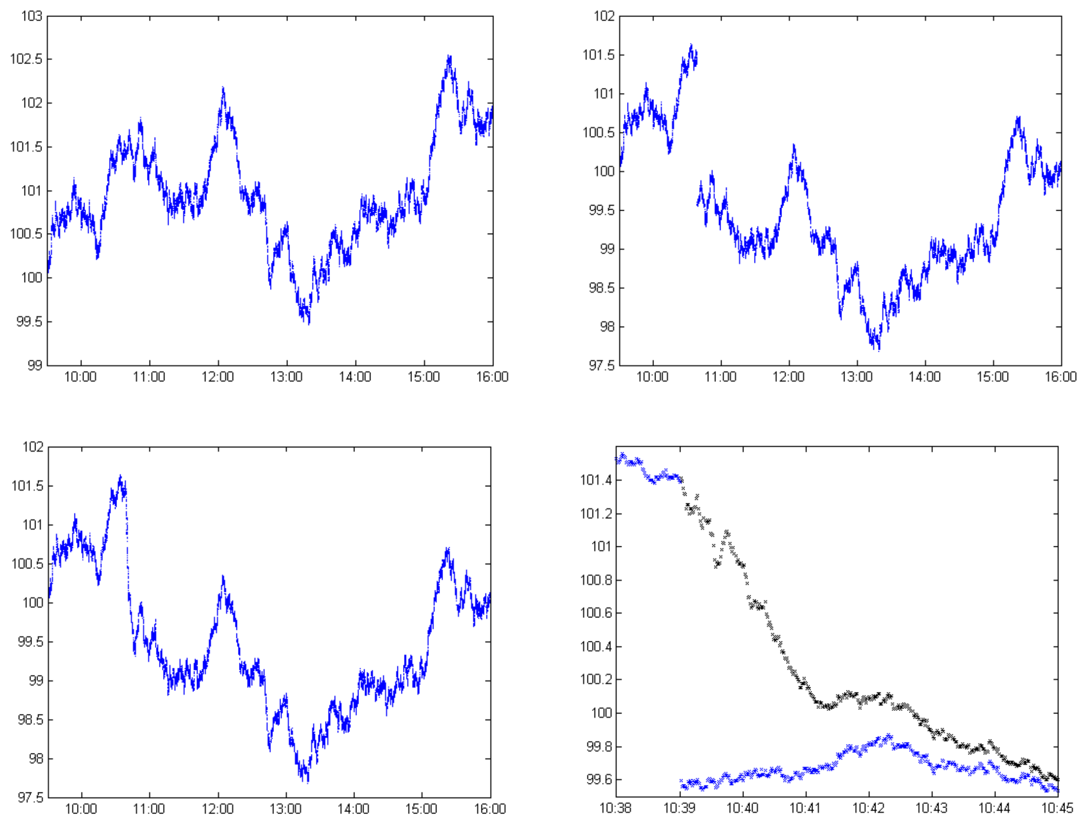


BNS, $t_0 = 250s$

Note: Figures illustrate how the AJ and BNS test results arise in the empirical case study for $\delta = 5$ minutes when increments are sampled from two different starting points. For the AJ test in first row, the bars are respectively $\left|y_i^\delta\right|^p$ and $\left|y_i^\delta + y_{i+1}^\delta\right|^p$, while the dotted lines are the accumulated values. The $\hat{S}$ statistic, the ratio of the second graph to first in each panel, is thus almost 8 for $t_0 = 170$, and about 1 for $t_0 = 550$. These values are also seen in figure 4. For the BNS test in second row, the bars are respectively $\left(y_i^\delta\right)^2$ and $\mu_1^{-2}\left|y_i^\delta\right|\left|y_{i+1}^\delta\right|$, and again dotted lines are accumulated values. The $\hat{R}$ statistic is similarly the ratio of the second graph to first in each panel. It is about 1 for $t_0 = 170$, and about 0.6 for $t_0 = 250$, values that also are seen in figure 4. These tests therefore show evidence of jumps for starting values in the second column, while for starting values in the first column, the statistics respectively point to continuity and fall far from asymptotic values under either hypothesis.

Figure 8:

Examples of Simulated Paths



Note: The figures show examples of simulated price series in levels for different models. In the top row to the left is the model under H_0, (26), while the figure to the right is from the same simulation but with a jump included according to the model (27). The bottom left figure adds noise around the jump following the gradual price adjustment model (25). The final figure zooms in on observations around the jump and illustrates how the price change enters immediately in the unobserved underlying price, but gradually in the observed price at this high observation frequency.

Research Papers 2008

2008-56: Carlos Capistrán and Allan Timmermann: Disagreement and Biases in Inflation Expectations

2008-57: Almut E. D. Veraart: Impact of time–inhomogeneous jumps and leverage type effects on returns and realised variances

2008-58: Dennis Kristensen and Yongseok Shin: Estimation of Dynamic Models with Nonparametric Simulated Maximum Likelihood

2008-59: Per Frederiksen and Frank S. Nielsen: Testing for long memory in potentially nonstationary perturbed fractional processes

2008-60: Thomas Q. Pedersen: Intertemporal Asset Allocation with Habit Formation in Preferences: An Approximate Analytical Solution

2008-61: Jean Jacod, Mark Podolskij and Mathias Vetter: Limit theorems for moving averages of discretized processes plus noise

2008-62: Giuseppe Cavaliere, David I. Harvey, Stephen J. Leybourne and A.M. Robert Taylor: Testing for Unit Roots in the Presence of a Possible Break in Trend and Non-Stationary Volatility

2008-63: Ole E. Barndorff-Nielsen, Peter Reinhard Hansen, Asger Lunde and Neil Shephard: Multivariate realised kernels: consistent positive semi-definite estimators of the covariation of equity prices with noise and non-synchronous trading

2009-01: Roman Frydman, Michael D. Goldberg, Søren JohanseN and Katarina Juselius: A Resolution of the Purchasing Power Parity Puzzle: Imperfect Knowledge and Long Swings

2009-02: Morten Ørregaard Nielsen: Nonparametric Cointegration Analysis of Fractional Systems With Unknown Integration Orders

2009-03: Andrés González, Kirstin Hubrich and Timo Teräsvirta: Forecasting inflation with gradual regime shifts and exogenous information

2009-4: Theis Lange: First and second order non-linear cointegration models

2009-5: Tim Bollerslev, Natalia Sizova and George Tauchen: Volatility in Equilibrium: Asymmetries and Dynamic Dependencies

2009-6: Anders Tolver Jensen and Theis Lange: On IGARCH and convergence of the QMLE for misspecified GARCH models

2009-7: Jeroen V.K. Rombouts and Lars Stentoft: Bayesian Option Pricing Using Mixed Normal Heteroskedasticity Models

2009-8: Torben B. Rasmussen: Jump Testing and the Speed of Market Adjustment