Jumps and Information Flow in Financial Markets

Suzanne S. Lee *

Georgia Institute of Technology

*The original idea in this article was first put forth in parts of my curriculum paper and PhD dissertation, submitted to the Graduate School of Business at the University of Chicago. I am very grateful to Per Mykland, George Constantinides, Federico Bandi, Ruey Tsay, and Pietro Veronesi for their helpful advice, suggestions, and comments. I also thank Frank Yu and Amit Goyal for their suggestions, Yuesong Li for his excellent research assistance, and the participants of the 2007 Volatility Conference at the University of Chicago and the 8th SAET Conference on Current Trend in Economics for their comments. Financial support for this research from Oscar Mayer Dissertation Fellowships, the National Science Foundation (DMS-02-04639), and the Financial Mathematics Program at the University of Chicago is gratefully acknowledged. Comments are welcome and errors are mine. Please address any correspondence to: Suzanne S. Lee, address: Georgia Institute of Technology, College of Management, Atlanta, GA 30332, phone: 404.822.1552, fax: 404.894.6030, e-mail: suzanne.lee@mgt.gatech.edu.

Jumps and Information Flow in Financial Markets

Abstract

We investigate the dynamics and predictability of stochastic jump arrivals in asset prices. We introduce a new two-stage, semi-parametric jump predictor test to determine informational covariates that affect jump occurrences up to the intraday levels. We find that jumps occur irregularly and jump size distributions are highly skewed to the left and have high excess kurtosis. Jumps tend to occur when there are stock-specific news releases and when analysts update their recommendations. We also show evidence of jump clustering in normal trading: the likelihood of jump arrivals becomes higher when there are jumps (especially, negative jumps) in previous trading hours.

JEL classification: G10, C14

Key words: jumps, predictor tests, information, jump clustering, partial likelihood

The predictability of asset return characteristics and the impact of information flow on asset prices have been two of the most enduring questions in financial economics and econometrics. There is a large stream of literature devoted to the predictability of the first and second moments of asset returns in relation to market information, using simple regression models and various types of time-varying variance models such as ARCH, GARCH, stochastic volatility, and forward-looking implied volatility models.[1] In contrast to the assumptions used in the aforementioned studies, recent studies widely recognize evidence of jumps that better explain higher moments of asset return distributions, such as excess kurtosis and skewness, and find an impact on financial management that is significantly different from impact of the diffusive innovation.[2] Our goal in this paper is to improve the understanding of the dynamics and predictability of such jump arrivals based on market information. Specifically, we introduce a new jump predictor test (JPT) to determine informational covariates that affect jump occurrence up to the intraday levels and provide empirical evidence in the U.S. equity markets.

Given the importance of jump dynamics, there is a recent literature on distinguishing jump risk from volatility risk in asset returns using discrete observations. [see Aït-Sahalia (2004), Huang and Tauchen (2005), Barndorff-Nielsen and Shephard (2005), Lee and Mykland (2006), and the references therein.] While this literature is focusing on testing the presence of jumps, there is no formal econometric test to further determine the effect of market information on prediction

[1] See Keim and Stambaugh (1986), Fama and French (1988 and 1990), among others, for prediction of the expected returns and Bollerslev, Chou, and Kroner (1992) with an excellent survey on the ARCH/GARCH literature for prediction of the second moments. The relation between stochastic volatility and information flow is investigated by Andersen (1996). The impact of mean return predictability on option pricing was first studied by Lo and Wang (1995) and many investigators such as Heston (1993), Hull and White (1987), and Stein and Stein (1991) study the impact of stochastic volatility but without jumps.

[2] Bates (1996), Bakshi, Cao, and Chen (1997), Aït-Sahalia (2002), Andersen, Benzoni, and Lund (2002), Pan (2002), and Carr and Wu (2006), among others, have shown the presence of jumps and their impact in option pricing and other financial applications.

of jump arrivals, which is one of the fundamental questions that should be addressed in financial decision making. Accordingly, we set out first to propose a new empirical test to determine which informational covariates do significantly affect jump arrivals in security returns so that the predictability of jump risk can be investigated, allowing the development of dynamic models for irregular jump risks separately from the usual volatility risk.

The second motivation of this study is for a better understanding of jump dynamics in individual equity markets, using our new empirical test. With high frequency stock returns, we can detect the date and time (within a day) of jump arrivals and their associated jump sizes, and find jump predictors of relevant information. Hence, we set as our second goal the thorough study of how jump risk systems in individual equity returns work. We specifically investigate in this paper when jumps occur more often, the size distribution of those detected jumps, and which information does trigger jump arrivals, among others. We choose to prove intra-day evidence, since the jump predictor test becomes more precise with high frequency observation. Therefore, we select candidate predictors from intra-day market information available. This part is also to illustrate how to apply our new proposed test to other markets.

Our methodology is summarized as follows. The new JPT is a two-stage semi-parametric procedure. Assuming the existence of both diffusive and jump risks under a continuous-time asset pricing model, we suggest first the application of the nonparametric jump detection test of Lee and Mykland (2006) ("Stage I"). This decomposes the two different types of risk in return distributions. The evidence is robust to model specification and under general assumptions such as stochastic volatility. It has been shown therein that this test detects the timing of jump arrivals with higher precision than other existing nonparametric tests. After the Stage I, we propose ("Stage II") the application of the test and estimation based on the maximum *partial* likelihood inference. Realizing the limitation in observing very small sized jumps because we only have discrete observations, we define it as a censoring problem that is often faced in the survival analysis literature. We prove that identifying market information related to jump arrivals will not be strongly affected by this censoring problem.

In order to detect the effect of any covariates on jump arrivals, econometricians choose the functional relationship between the covariates and the jump arrivals in Stage II. Because Stage II is in essence a novel part of this paper, we explain carefully the theory and approximation of the likelihood function to make an inference, and directly provide the asymptotic distribution of our JPT and its related tests. Since the results in both stages are based on asymptotic arguments, we investigate the finite sample performance of the test, using Monte Carlo simulation studies. We find that performance is good so long as we use high frequency observations such as 15-minute returns. It broadly contributes to various literature by providing an efficient empirical method to solve similar problems. For example, it can be applied not only to individual stock prices but also to all kinds of other financial time series such as bond prices, exchange rates, interest rates, volatility, etc. in relation to market information, so long as target returns and chosen information covariates are available at the same observation frequencies.

After we introduce JPT, we apply it to all the individual component stocks of the Dow Jones Industrial Average (DJIA) index, collected from the Trade and Quote (TAQ) database. We select the actively traded stocks that are more likely to be liquid. The empirical results appear to confirm evidence of irregular jump arrivals in a larger class of individual equities transacted on the New York Stock Exchange (NYSE). Stage I of JPT indicates that the jump size distributions for none of our sample individual equities are normal; they are skewed to the left and have high excess kurtosis. Another finding in Stage I suggests that if there are jumps on a trading day, they tend to come in the morning (about 90% before 10:30 am). However, average number of jumps a year for each company is about 20, much less than total number of trading days a year. This shows that the trading mechanism with interrupted overnight market itself does not induce jumps and we are motivated to study which information triggers such jump arrivals.

In Stage II, we first set the information jump predictors to be indicators of trading hours near pre-scheduled earnings announcement dates (EADs)[3] or a company-specific news release using COMPUSTAT and I/B/E/S database with confirmations of a *Factiva* news search. We find significant evidence of jumps occurring during morning hours on the EADs or next days if the announcement is released after trading hours on previous days. Secondly, using *First Call Recommendation History* database, we test if analysts' recommendation updates can affect jump arrivals, and find evidence of jumps immediately or at the earliest possible time after analysts release their recommendations. Finally, we present evidence of jump clustering in normal trading by showing that jumps in previous trading hours have a much stronger positive impact on the likelihood of jump arrivals in the near future such as next day. Interestingly, we observe previous negative jumps increase the likelihood of future jumps more than positive jumps. In other words, uncertainty increases more after negative jumps, hence there is potential for market timing. Since our method allows the analysis of directions of jumps that are distinguished from volatility risk, we also examine the risk mechanism which can not be addressed only by volatility clustering. We expect that evidence of jump predictability shown here should direct us to a different financial management, depending on jump timing and other market information.

The remainder of the paper is organized as follows. Section 1 sets up a theoretical framework to detect jumps and to test the effect of information variables on irregular jump arrivals. Section

[3]More precisely, the indicators are set for trading hours on pre-scheduled announcement dates and next days to capture news release after trading hours.

2 introduces the two-stage, semi-parametric JPT and presents finite sample performance along with simulation results. After we empirically examine the characteristics of jump dynamics in DJIA component stock returns and the possible predictability of jumps in Section 3, we discuss the implications of identifying the dynamics of stochastic jump arrivals in Section 4. Finally, we conclude in Section 5 with a summary of our findings.

1 A Theoretical Model

We employ a one-dimensional asset return process with a complete probability space $(\Omega, \mathcal{F}_t, \mathcal{P})$, where $\{\mathcal{F}_t : t \in [0, T]\}$ is an information filtration for market participants up to time t, and $\mathcal{P}$ is a data-generating measure. Let the continuously compounded return be written as $d \log S(t)$ for $t \geq 0$, where $S(t)$ is the asset price at t under $\mathcal{P}$. For simplicity, we illustrate here a univariate model of individual asset returns. Extension to a multivariate model of cross-sectional returns is straightforward with a correlation structure between risks from different assets in a market, and is studied in Lee (2007).

According to empirical evidence of jumps and relevant models from the literature, as discussed in the introduction, we assume a return process with diffusive risk and irregular jump risk such that asset price $S(t)$ is represented by the following stochastic differential equation (SDE):

$$d \log S(t) = \mu(t)dt + \sigma(t)dW(t) + Y(t)dJ(t), \tag{1}$$

where $W(t)$ is a $\mathcal{F}_t$-adapted standard Brownian motion and the drift $\mu(t)$ and spot volatility $\sigma(t)$ are $\mathcal{F}_t$-adapted processes. Assumptions on the drift and volatility are the same as in Lee and Mykland (2006), and is restated in the Appendix for reader convenience. These assumptions allow time-varying drift and stochastic volatility, each of which can depend on itself. This model without its jump component describes diffusive risk in returns due to normal randomness in markets. So long as drift and volatility satisfy the assumptions, we do not impose any further specifications on them.

For the jump risk, we include $Y(t)dJ(t)$, which describes more dramatic, irregular, and unusually large risk in returns and imposes a parametric structure as described in Subsection 1.1 below. In essence, this allows a parametric specification to incorporate time-varying information flow into jump intensity by means of a general function, which is the novel difference between the assumption presented here and those existing in the literature. Except for the intensity structure, all the other part of model remain nonparametric.

1.1 A Sub-Model for Irregular Jumps

Early research based on jump diffusion models assumed the rate of jump arrival (jump intensity) to be constant so that jump occurs regularly. However, recent papers by Andersen, Benzoni, and Lund (2002), Chernov, Gallant, Ghysels, and Tauchen (2003), Pan (2002), Johannes, Kumar, and Polson (1999), and Maheu and McCurdy (2004), among others, recognize the fact that the process governing the jump arrival can be dynamic and heterogeneous with respect to the type of news and other relevant market information. Accordingly, we frame a general structure describing the jump arrival mechanism that incorporates heterogeneous information flow over time.

We set $J(t)$ to be a non-homogeneous Poisson-type counting process with stochastic intensity $\Lambda_\theta(t)$. This time-varying cumulative intensity process $\Lambda_\theta(t)$ is specified by a q-dimensional parameter $\theta = (\theta_1, .., \theta_q) \in \Theta$, an open subset of the q-dimensional Euclidean space. Thus, we write

$$\Lambda_\theta(t) = \Lambda(X(t); \theta) = \gamma(t, X(t); \theta), \tag{2}$$

where $X(t)$ denotes the time-varying informational covariates that can affect the likelihood of jump arrival, and γ is a general function of time and the covariates. Assumptions on this general function γ are stated in the Appendix. Essentially, the integrated intensity function is required to be continuous and differentiable so that the martingale central limit theorem can hold and the solution for the corresponding score function exists and is consistent. $Y(t)$ represents jump size, a predictable process. [see Protter (2004) for definition.] Its mean $\mu_y(t)$ and standard deviation $\sigma_y(t)$ are also $\mathcal{F}_t$-predictable processes. We assume that jump sizes $Y(t)$ are independent of each other and identically distributed and independent of other random components. Hence, the jump counting $J(t)$, the diffusion $W(t)$, and the jump size $Y(t)$ are all independent from one another.

2 Discrete Observations and Overall Inference Methods

In this section, we explain how our overall test and estimation are accomplished. We focus our attention on the analysis of jump risk that is less persistent and more irregular, hence harder to predict. We briefly mention applicable methods for more persistent stochastic volatility inference. This order of discussion does not necessarily have to coincide with the order of actual analysis of the model because the suggested inference method for each risk is robust to the presence of each other risk. Since this paper focuses more on irregular jump analysis, we concentrate on its estimation.

Although our model assumes continuous information flow and hence, continuous asset price movement, only discrete observations are applicable to our econometric test and estimation of model. This has made difficult the decomposition of different types of risk, diffusive and jump risk; investors tend to prefer separate management strategies with regard to these risk types depending on their level of risk aversion. To resolve this issue, we apply a recently developed, nonparametric test by Lee and Mykland (2006) to distinguish jump arrivals and their sizes.[4] This

[4]For related studies, see Barndorff-Nielsen and Shephard (2006), Jiang and Oomen (2005), and Aït-Sahalia and

test decomposes two risks so that separate analysis for the component risks in returns are possible.

Building on such decomposition, we propose using a new approach for the jump risk analysis, a two-stage, semi-parametric JPT described below. The idea underlying this JPT is simple. We first extract jumps from the return series over time by applying the multiple nonparametric jump tests of Lee and Mykland (2006). The detected jumps become partial observations of the continuous-time jump counting process $J(t)$. Using these detected jump observations, we apply the *maximum partial likelihood* tests to determine the relevant information covariates. We graphically illustrate the idea of two stages in panels A and B of Figure 1. We explain each step in the following subsections and provide the limiting distribution of the test statistic to use as a benchmark.

For the analysis of diffusion risk, $\sigma(t)$, it has been proven that we can consistently estimate the instantaneous stochastic volatility by the realized bi-power (more generally, multi-power) variation, even if there are jumps in a market, so long as jumps in that market are rare (more precisely, so long as they are not infinite-activity jumps), and high-frequency observations are available. Therefore, application of the realized multi-power variation is sufficient for the separate analysis of the diffusive risk part of our model.[see Aït-Sahalia (2004), Barndorff-Nielsen and Shephard (2004), Barndorff-Nielsen, Shephard, and Winkel (2005), and Aït-Sahalia and Jacod (2006b) for more detail.]

We assume a time horizon T and a number of observations n. The observation of asset prices $S(t)$ and informational variables $X(t)$ occurs only at discrete times $0 = t_0 < t_1 < ... < t_n = T$. For simplicity, we set observation times as equally spaced: $\Delta t = t_i - t_{i-1} = \frac{T}{n}$. This simplified assumption can easily be generalized to non-equidistant cases by letting $\max_i(t_i - t_{i-1}) \to 0$.

Jacod (2006a).

2.1 Two-Stage Semi-Parametric JPT

In this section, we provide a detailed discussion of the JPT. The purpose of this test is to investigate the effect of any informational covariates on irregular jump arrivals and to determine whether they are significant or not. We call this JPT *two-stage* and *semi-parametric* because it combines two different procedures. In Subsection 2.1.1, we discuss the first stage of the JPT, the application of nonparametric test of Lee and Mykland (2006). We discuss the second stage analysis of applying maximum partial likelihood predictor tests in Subsection 2.2.2.

2.1.1 Stage I: Lee and Mykland (2006)'s Nonparametric Jump Test

We explain the application of the jump test to high-frequency time series of returns over time to extract jump series in the first stage. The idea behind the jump test by Lee and Mykland (2006) is to compare the realized return at every observation time to a consistently estimated instantaneous volatility, $\sigma(t)$, using observations prior to the testing time. Mathematically, the statistic $\mathcal{L}(i)$, which tests at time t_i whether there was a jump from t_{i-1} to t_i, is defined as

$$\mathcal{L}(i) \equiv \frac{\log S(t_i)/S(t_{i-1})}{\widehat{\sigma(t_i)}}, \tag{3}$$

where

$$\widehat{\sigma(t_i)}^2 \equiv \frac{1}{K-2} \sum_{j=i-K+2}^{i-1} |\log S(t_j)/S(t_{j-1})||\log S(t_{j-1})/S(t_{j-2})|. \tag{4}$$

K in the definition is a window size within which a local movement of the process is considered. With the first $K-1$ realized returns in the window, the *instantaneous volatility* is estimated as in equation (4) by the *realized bipower variation.* The last realized return from t_{i-1} to t_i in the window is then compared by taking the ratio of this estimated volatility to the last return in order to determine whether a jump arrived between t_{i-1} and t_i, and how large that jump size was. For example, if Δt = 5 minutes, t_i =10:05am, and K = 10, then we test for a jump by examining the relative magnitude of a realized return from 10:00am to 10:05am compared to instantaneous volatility estimated using 5-minute returns from 9:15am to 10:00am. The volatility estimator used in the test is robust to the presence of jumps in the model. The optimal window size, K, is specifically recommended in Lee and Mykland (2006), depending on the observation frequency.

With the discrete observations from our continuous-time model as set up in Section 1, all discrete times, t_i for all $i \in \{1, 2, .., n\}$, can be categorized into two sets of times over the whole time horizon. One set of times contains i's such that the interval $(t_{i-1}, t_i]$ includes no jumps and we call it "No Jump Times" (NJT_n). The other set of times contains i's such that the interval $(t_{i-1}, t_i]$ does include a jump and we call it "Jump Times"(JT_n). Lee and Mykland (2006) prove that the test statistic $\mathcal{L}(i)$ with $i \in \mathrm{NJT}_n$ approximately follows a normal distribution, with its mean zero and variance $\sqrt{2/\pi} \approx 0.7979$, as the frequency of observation is increased. On the other hand, the statistic $\mathcal{L}(i)$ with $i \in \mathrm{JT}_n$ converges to ∞ as we increase the frequency of observation. Combining these results with the extreme value theory by Galambos (1978) suggests the jump detection rules, based on the asymptotic distribution of maximums of the statistics for $i \in \mathrm{NJT}_n$. We use the critical values from this distribution as recommended to distinguish the random jump arrivals.

So long as we use high frequency observations, application of the jump test makes negligible the probability of misclassification of jumps due to either our failing to detect actual jumps or to spuriously detecting jumps when there is actually no jump at any given testing time. This proposition is also proved in multiple tests, and is shown to hold for jumps with several different jump sizes. We assume that when we observe a jump in an interval, the order of magnitude of the jump in the realized returns dominates that of the diffusion part of the model. Hence, we use the return itself as the jump size, which can also be one of the information covariates in Stage II.

2.1.2 Stage II: Maximum Partial Likelihood Estimation and Tests

In this subsection, we explain that we can use "usual" maximum likelihood estimation and related tests on jumps detected in Stage I, in order to investigate the effect of informational covariates on jump arrivals. Although our irregular jump arrival sub-model is specified by the "continuous-time" jump counting process, again, inference should depend on discrete observations. The best possible way to observe data for the jump counting process is by using those detected from the nonparametric test at Stage I. We demonstrate that such an approach is sufficient for our purpose. We first show that the likelihood function for the continuous-time Poisson processes of our sub-model can be approximated by partitioning it into a discrete counterpart. Then, observations for this approximate likelihood are comprised of detected jumps in Stage I. We apply in this stage the partial likelihood results based on the martingale central limit theorem. The mathematically rigorous regularity conditions under which the results hold are stated in the Appendix, and related proofs can be found therein.

To demonstrate the sufficiency in this particular problem, we first discuss background mathematics. Remember, we have n observations between $[0, T]$, with $t_0 = 0 < t_1 < t_2 < ... < t_n = T$. We define the product-integration $\widetilde{\prod}$ over $[0, T]$ of any *cadlag* (left continuous and right limit) function with $t_i \in [0, T]$ as

$$\widetilde{\prod}_{s\in[0,T]}(c_1 + c_2 dg(s))^{c_3 + c_4 dh(s)} = \lim_{\Delta t \to 0} \prod_{1 \le i \le n} (c_1 + c_2 dg(t_i))^{c_3 + c_4 dh(t_i)}, \tag{5}$$

where c_1, c_2, c_3, and c_4 are some constants, $dg(t_i) = g(t_i) - g(t_{i-1})$, $dh(t_i) = h(t_i) - h(t_{i-1})$, and $t_0 = 0 < t_1 < t_2 < ... < t_n = T$ are discrete times to make a partition of $[0, T]$. Although this equal-distance assumption again can be generalized by letting $\max |t_i - t_{i-1}| \to 0$, we set them constant in this paper.

For our irregular jump counting sub-model, we have the well-defined continuous-time likelihood function $\widetilde{L}$ as

$$L(\widetilde{\theta|\mathcal{F}_t}) = \widetilde{\prod}_{s\in[0,t]} d\Lambda_\theta(s)^{dJ(s)} \widetilde{\prod}_{s\in[0,t]} (1 - d\Lambda_\theta(s))^{1-dJ(s)}. \quad (6)$$

The above definition of product-integration and the likelihood function suggest that we can approximate the likelihood function by replacing the instantaneous changes by the increments of $J(t)$ and $\Lambda_\theta(t)$ over intervals from t_{i-1} to t_i, and forming the corresponding finite products. Hence, the actual data analysis can be done by the approximate full likelihood function L, such that

$$L(\theta|\mathcal{F}_T) = \prod_{1\le i\le n} d\Lambda_\theta(t_i)^{dJ(t_i)} \prod_{1\le i\le n} (1 - d\Lambda_\theta(t_i))^{1-dJ(t_i)}, \quad (7)$$

so long as there is no error in observing jumps. However, since we depend on test results from Stage I using discrete data that are mixed with diffusion, there can be possible misclassification of jumps in our Stage I, hence there can be errors in observing jumps, especially when jump sizes are extremely small. Although it has never been recognized, this can also be considered as a form of censoring problem due to discrete observations. In the survival analysis literature, when data cannot be observed beyond a certain level in their distribution, they are called *right censored* or *left censored* problems, depending on which part of the data distributions are unobservable. We solve this issue in a fashion similar to survival analysis but, in our case, we call it *middle censored*, because very small sized jumps tend not to be detected if we do not use continuous observations, which is likely to be the case in this type of analysis.

Here we set a partial likelihood function using the detected jumps from Stage I as follows (this is our objective function to be maximized):

$$PL(\theta|\mathcal{F}_T) = \prod_{1\le i\le n} d\hat{\Lambda}_\theta(t_i)^{d\hat{J}(t_i)} \prod_{1\le i\le n} (1 - d\hat{\Lambda}_\theta(t_i))^{1-d\hat{J}(t_i)}, \quad (8)$$

where $\hat{\Lambda}(t_i)$ and $\hat{J}(t_i)$ are the filtered partial observations which may include errors from Stage I.

Now, in Proposition 1, we demonstrate that this partial likelihood based on the detected jumps in Stage I is sufficient for making inferences on the effects of $X(t)$ on $\Lambda_\theta(t)$.

Proposition 1: Partial Likelihood is Sufficient.

Suppose that Assumption C in Appendix holds. Let $L(\widetilde{\theta|\mathcal{F}_T})$, $L(\theta|\mathcal{F}_T)$, and $PL(\theta|\mathcal{F}_T)$ be defined as in equation (6), (7), and (8). Then,

$$P(L(\theta|\mathcal{F}_T) \neq PL(\theta|\mathcal{F}_T)|N) \to 0 \text{ as } \Delta t \to 0, \tag{9}$$

Furthermore, with probability 1, both $PL(\theta|\mathcal{F}_T)|N)$ and $L(\theta|\mathcal{F}_T)$ converge to the same limit: $L(\widetilde{\theta|\mathcal{F}_T})$ as $\Delta t \to 0$.

We now apply the martingale central limit theorem to obtain the following results. Hereafter, we only use the partial likelihood and omit the hat notation. The log-partial approximate likelihood function, ignoring the constant term, becomes

$$\mathcal{C}(\theta) = \log PL(\theta) = \sum_{1\le i\le n} \log d\Lambda_\theta(t_i) dJ(t_i) + \sum_{1\le i\le n} \log(1 - d\Lambda_\theta(t_i))(1 - dJ(t_i)). \tag{10}$$

To obtain the partial approximate score function, we differentiate the above function with respect to θ. Assuming that we may interchange the order of differentiation and summation, the vector $\mathcal{U}(\theta)$ of partial approximate score statistics $\mathcal{U}^j(\theta), j = 1, .., q$, is

$$\mathcal{U}^j(\theta) = \sum_{1\le i\le n} \frac{\partial}{\partial\theta_j} \log d\Lambda_\theta(t_i) dJ(t_i) + \sum_{1\le i\le n} \frac{\partial}{\partial\theta_j} \log(1 - d\Lambda_\theta(t_i))(1 - dJ(t_i)) \tag{11}$$

With probability tending to 1, it has been shown that there exists a consistent solution $\hat{\theta}$ of the equation $\mathcal{U}(\theta) = 0$, and that the solution is asymptotically normally distributed around the true parameter value, with a covariance matrix, that is, usual maximum likelihood results. The covariance matrix can be estimated in the usual manner by $\mathcal{Z}^{-1}(\hat{\theta})$, where $-\mathcal{Z}(\theta)$ is the matrix of second-order partial derivatives of the log-partial likelihood function. Thus, we have

$$\mathcal{Z}(\theta) = -\sum_{1\le i\le n} \frac{\partial^2}{\partial\theta_j\theta_l} \log d\Lambda_\theta(t_i) dJ(t_i) - \sum_{1\le i\le n} \frac{\partial^2}{\partial\theta_j\theta_l} \log(1 - d\Lambda_\theta(t_i))(1 - dJ(t_i)). \tag{12}$$

Proposition 2 below lists all related predictor tests for the purpose of our study.

Proposition 2: Jump Predictor Test (JPT)

Suppose that Assumptions C and D in Appendix hold. Let θ be the q dimensional effect parameter of q information covariates that affect jump arrivals, and let $\hat{\theta}$ be the maximum likelihood estimate based on the partial likelihood function as in equation (8). Under the simple null hypothesis $H_0 : \theta = \theta_0$, as long as $T \to \infty$ and $\Delta t \to 0$,[5] there is a consistent estimator $\hat{\theta}$, which is a solution to $\mathcal{U}(\theta) = 0$, and is asymptotically normal around its true value θ_0. And our JPT statistic, defined as

$$(\hat{\theta} - \theta_0)' \mathcal{Z}(\hat{\theta})(\hat{\theta} - \theta_0) \tag{13}$$

converges in distribution to a chi-square random variable with degrees of freedom q, $\mathcal{X}^2(q)$.[6]

2.2 Simulation Study

In this subsection, we examine the effectiveness of the jump predictor test using a Monte Carlo simulation. The simulation study shows the finite sample performance of the asymptotic results. Overall simulation results prove that the JPTs do perform well in distinguishing the effects of

[5]Since $\Delta t = \frac{T}{n}$, as long as $n \to \infty$ faster than $T \to \infty$, the results hold.

[6]Also the approximate score test statistic $\mathcal{U}(\theta_0)' \mathcal{Z}^{-1}(\theta_0)\mathcal{U}(\theta_0)$ and the partial likelihood ratio test statistic $2\left(\mathcal{C}(\hat{\theta}) - \mathcal{C}(\theta_0)\right)$ have the same asymptotic distribution as the Wald test statistic, as in equation (13).

information covariates. For series generation, we use the Euler-Maruyama Stochastic Differential Equation (SDE) discretization scheme [see Kloeden and Platen (1992)], an explicit order 0.5 strong and order 1.0 weak scheme. To avoid the starting value effect on series generation, we discard the burn-in period.

We consider the simplest case of our general model as stated in Section 1, a constant volatility $\sigma = 0.3$ and an information covariate to affect the stochastic jump intensity in addition to a constant intensity. We set $dX(t)$ to be an indicator of monthly information release in this study, which becomes equal to 1 once every month (12 times a year) and equals 0 otherwise. We assume the intensity is an affine function of the information covariate. Hence, $d\Lambda_\theta(t) = \theta_0 + \theta_1 dX(t)$. We suppose an analyst tests whether the monthly information release is a significant jump predictor by the proposed test. Three hundred series of returns over 1 year are simulated at an intraday 15-minute frequency. The usage of 15-minute returns is recommended for reasonable results from Stage I. [see Lee and Mykland (2006)]. θ_0 and θ_1 are set at 0.01 and 0.99 so that the expected value of the jump intensity over a year is about 5%. The mean and standard deviation of jump size distribution are assumed at 0 and $3 \times \sigma$ respectively in Table 1, which shows a representative example of the simulation results. We employ an optimization procedure to solve the non-linear system of equations of parameters. The standard errors and 95% confidence intervals reported in the table are based on the asymptotic normal distribution of the parameter values stated in Proposition 2. Table 2 shows the averaged results of our simulation studies with three different levels ($1 \times \sigma$, $2 \times \sigma$, and $3 \times \sigma$) of standard deviation of jump size distributions. The results in Table 2 indicate that in the finite samples, the jump size change does not strongly affect the conclusion reached by the test on the effect of informational covariates.

3 Empirical Analysis for U.S. Equity Markets

In this section, we perform data analysis using our new jump predictor test to investigate the effect of informational covariates on stochastic jump arrivals in the major U.S. individual equity markets. Both stages of the JPT are based on asymptotic theories which assume that the time distance between two discrete observations is small. Although our goal is to illustrate the final effect of chosen predictors on jump arrivals, the intermediate empirical evidences in Stage I are also informative and important and have never been presented in the literature. Hence, we separately report the results for Stage I and Stage II in Subsections 3.2 and 3.3, respectively.

3.1 Data

We select the most actively traded U.S. large-cap equities of the DJIA that are traded on the NYSE. Data are collected from the TAQ database. The TAQ database contains continuously recorded trading information such as transaction times, prices, exchange, and volume information beginning with the year 1993. Although an analysis using the whole data set beginning with the year 1993 is, in principle, possible, we choose more recent data that cover a long enough period to achieve our objective. Our selected sample extends from July 1, 2001 to June 30, 2006, for a total of 1,256 trading days including all available records from 9:30am to 4:00pm; this represents the most recent data set available, and it has never been investigated in the literature.[7] We select only DJIA firms that are traded on NYSE in order to maintain enough degrees of liquidity and to maintain similar organization of trading mechanisms and trading hours across stocks. For this

[7] The sample period of 5 years was chosen because we use high frequency transaction data to increase the precision of the test in Stage I and because 5 years is reasonably long enough for Stage II to give reliable empirical results. [Also see Andersen, Bollerslev, Diebold, and Ebens (2001), who also used a 5-year sample period for their high frequency data analysis.]

reason, two among 30 stocks are excluded because they are traded on the National Association of Security Dealers Automated Quotation system (NASDAQ).

Since the simulation study shows that a 15-minute frequency is high enough for our test to achieve a sufficient power for our JPT, we choose to use 15-minute stock returns by taking the differences of log transaction prices at 15-minute intervals and multiplying all returns by 100 to present them as percentages. A 15-minute frequency is long enough for our empirical results not to be greatly affected by the presence of market microstructure noise. When dealing with high-frequency intraday returns, as noted in Andersen, Bollerslev, Diebold, and Ebens (2001), market microstructure effects such as bid-ask bounce can be avoided with 5-minute or longer frequency. Therefore, the evidence presented in this paper is robust to the effect of market microstructure noise.

To avoid unnecessary data recording errors, we also preprocess the raw data as follows. All stocks we select from the DJIA index in this study are assured to pass the active trade filter (50 trades per day[8]). For transactions that happen at the same time, we take the first transaction price recorded in the database. We exclude all recording errors such as zero prices. As noted in Aït-Sahalia, Mykland, and Zhang (2006), *bounce-back* type data errors are caused by extreme round trips of recorded prices to unreasonably different price levels. If returns are followed by a return with opposite signs and the same magnitudes and if the magnitudes of jumps are significantly different from those without the bounce-back effect, they are removed. However, this can be partly due to microstructure noise and, again it is rare to have such type of errors at 15-minute returns.

In Stage II, we test whether earnings announcement, analysts' recommendation, or previous

[8]This filtering rule was used in Easley, O'Hara, and Srinivas (1998), Chan, Chung, and Fong (2002), and Tookes (2006) for their high frequency data analysis.

jump arrival can be the predictors of future jump arrivals. We discuss sample selection of information covariates and how we construct the jump predictors in corresponding subsection in more detail.

3.2 Empirical Results of Detecting Jumps: Stage I

In this subsection, results from Stage I are presented in Tables 3, 4, and 5. The significance level for the first stage nonparametric jump detection test is 5%. The outcomes of Stage I are each jump size, jump arrival date, and time. We do not assume that there is only one jump per day, but we do assume that when a jump occurs, the jump size dominates the return because the order of jump size magnitudes dominates the order of diffusion magnitudes, as mentioned in Subsection 2.1.1.

Table 3 shows the summary statistics that characterize the jump size distribution of the individual component stocks in the DJIA, that is, the mean, median, standard deviation, skewness, kurtosis, lower and upper quartiles, and minimum absolute jump sizes. As a reference, the numbers of detected jumps for each company during the 5-year period are also included in the ReJ column. All the figures in Table 3, except the ReJ column, are percentages. The majority of stocks have a negative mean jump size, ranging from -1.264 for XOM to 0.564 for GM. The cross-sectional average of means is significantly different from zero. However, the cross-sectional average of medians of all stock jump sizes is positive, although it is not significantly different from zero. The standard deviations are around a mean of 3.968 with standard error 0.383. Skewness values are all negative, except for two stocks (IBM and T), with a significantly negative minimum of -8.635 for PG. For all but three stocks, the absolute magnitude of the negative minimum jump size is greater than that of the stock's positive maximum jump size. Finally, the minimum absolute jump size is always greater than 0.411, which indicates that very small sized jumps are middle censored, as discussed above, due to discrete observations such that they are not detected by the nonparametric test in Stage I. The jump size characteristics presented in Table 3 imply that all individual stocks in our sample have fat-tailed return distributions in physical data-generating measures. Frequent and larger-sized jumps in individual stock returns are likely to make the tails fatter.

Table 4 shows the average number of jumps detected in Stage I over 1 year, 1 month, and 1 day. Each year, stocks in the sample experience 22 jumps on average, from 13 for MMM to 29.4 for PFE. Every month, one to two jumps occur on average in each stock. The daily average rate of jump arrival is 0.0881. This rate is calculated with the assumption that the jump arrival rate is constant over time for summary purposes. We observe, however, that they do not occur regularly. Therefore, models with constant jump intensities are not appropriate. Table 5 presents more specifically when in a day these jumps arrive. It reports the percentages of detected jumps occurring during specific time intervals in a trading day among all realized jumps. We divide the trading hours of the NYSE, 9:30am to 4:00pm, into nine categories. We find that if a jump comes, more than 90% of jumps arrive before 10:30am, around the time of market opening. However, it is observed that opening prices do not always include jumps.

Summarizing Tables 4 and 5, we conclude that if jump occur, they tend to arrive in the morning, but every overnight return does not necessarily includes jumps: there are far fewer jumps than the number of trading days. The NYSE trading mechanism for opening markets provides a natural controlled experiment framework to study whether the market interruption itself is the cause of jumps or large deviations in stock prices. Without information waiting to be reflected in prices, the interruption itself does not trigger jumps. They are triggered when the demand for trading is increased due to the accumulated information flow but the market is not liquid under the interrupted market organization. We formally test this hypothesis in the next subsection. More specifically, we investigate which market information triggers more jumps.

3.3 Empirical Results of Identifying Jump Predictors: Stage II

3.3.1 Earnings Announcements

In this subsection, we test the hypothesis whether jumps are triggered by the stock specific scheduled earnings announcements for our selected large U.S. stocks. We collect earnings announcement dates for our sample period of 5 years for each of DJIA component companies from the COMPUSTAT database up to the year 2005 and I/B/E/S database for the year 2006 because COMPUSTAT covers earnings report dates only until 2005. To minimize the data error, release dates are compared between the COMPUSTAT and I/B/E/S databases and also confirmed with the Factiva real-time news database. If the dates from these sources are different, we use the timing information from Factiva due to possible recording errors, also mentioned in Dubinsky and Johannes (2006). Since earnings are often released after trading hours, in order to capture the effect during the first available time in which the information can be reflected into prices, we let the announcement dates be both the reported date and the next date in setting up the covariates for our analysis.

We perform the following tests on each stock. To learn if the news release triggers jump arrivals and when it is more likely to affect them, we create three different information release time indicators, denoted as $dX_0(t)$, of opening hours, morning hours, and all trading hours on earnings announcement dates for each stock. We assume the covariates are related to jump arrivals in an affine function: $d\Lambda_\theta(t) = \theta_0 + \theta_1 dX_0(t)$. To save space, we present only representative cases focusing on three individual stocks (C, DIS, and IBM) in Table 6.[9] As can be seen in Table 6,

[9]Complete unreported results for other companies can be requested from the author.

the magnitude of θ_1 becomes much larger for the opening hours on earnings announcement dates. The size of θ_1 for the morning effect is almost 10 times greater than that for the whole day effect on the EADs. The p-values for all three informational covariates are less than 5%, which leads us to conclude that all the covariates can be regarded as the predictor of jump arrivals. The degrees of significance can be taken as the degrees of precision of the jump predictor: if the p-values are lower, the precision is higher so long as they are lower than the chosen significance level. In this sense, the announcement day opening indicator is the most precise jump predictor. We find similar results for all the other companies and the same conclusion can be drawn from other cases. This evidence partially supports the setting of the individual equity option pricing model in Dubinsky and Johannes (2006), which incorporates deterministic jump events conditional on scheduled EADs, although we have many more jumps than four times that is the number of EADs a year. This finding is also consistent with the notion that significant financial market jumps are related to responses to (macroeconomic) news announcements as in Andersen, Bollerslev, Diebold, and Vega (2003).

3.3.2 Jump Clustering

We examine another question, that is, whether previous jump occurrences change the rate of future jump arrivals in normal trading hours. In short, we test jump clustering evidence. Literature use the term volatility clustering for the fact that volatility shows positive and significant autocorrelation.[10] We use here the notion of "clustering" similarly for our study. More precisely, jump clustering means jump arrivals tend to follow previous jump arrivals. Since our test distinguishes jumps from volatility, clustering evidence for jumps is also separate from volatility clustering phenomenon.

[10]Small changes in returns tend to follow small changes and large changes in returns tend to follow large changes.

Table 7 reports the p-values of the JPT statistics for all companies included in the DJIA without presenting the magnitudes of their coefficients. We find that more jumps in preceding trading hours have a stronger positive impact on the likelihood of future jump arrivals. To show this effect, we select two different indicators. One is an indicator of one jump occurrence within the previous 7 trading hours ($dX_1(t)$). The other is an indicator of two jump occurrences within the previous 14 trading hours ($dX_2(t)$). We again set the stochastic jump intensity to be an affine function of each of these indicators. As can be seen in the table, we find that the p-values for all but 4 stocks are less than 5% with the predictor of one jump occurrence during the previous 7 trading hours. We find much stronger results with the second predictor, two jump occurrences during the previous 14 trading hours. All the p-values are much less than 5%. Except for two stocks, they are less than 1%. Hence, this simply proves evidence of jump clustering in normal trading. The effect gets only stronger with more jumps.

We then question what kind of jumps trigger more jumps in the near future. For this, our predictors for jump clustering effect are categorized into two kinds: positive and negative jumps for each stock of our sample companies. As found in earlier part of this subsection, both of them are predictors but the degrees of predicting powers, measured in terms of p-values, are different. We find that for 73% of our sample companies, the likelihoods of jump occurrence after negative jumps are greater than that after positive jumps.[11] In other words, investors find equity market more uncertain after negative jumps than after positive jumps. This clustering evidence can be used as a market timer if one is interested in taking advantage of future jump risks.

[11]We also study if positive (negative) jumps are triggered more by positive jumps or negative jumps, which we can term as momentum and reversal effect in jump risks. We did not find consistent patterns across 26 different sample stocks except that negative jumps trigger more jumps (both positive and negative) than positive jumps.

3.3.3 Stock Analyst Recommendation

In this subsection, we investigate whether release of analyst recommendation, another form of information, can be a jump predictor. We obtain the comprehensive real-time analyst recommendation history of the DJIA companies during our sample period from 2001 to 2006. Intra-day recommendation data are collected from the First Call, a subsidiary of Thomson Corporation, which most brokerage firms and institutional investors depend on in order to disseminate their research reports electronically to their clients through a news wire service. It provides the exact dates and time-stamps of analyst recommendation updates, measured within 1 minute, which allow us to learn when the information becomes available to investors and whether it affects jump arrivals for our study. To reduce bias due to sample selection, we include all types of recommendations changes by all analysts reported in the database, unlike Womack (1996) who examines immediate market reactions to dramatic recommendation changes (added or removed to buy (sell) recommendation) made by the highest rated U.S. brokerage research departments. Hence, the analysis includes not only added to buy (sell) recommendation or removed from buy (sell) recommendations but also changes from buy to strong buy or sell and other kinds. Each recommendation record from the database contains the ticker symbol of the corresponding company, date and time of update (up to minutes), and one-to-five point recommendation scales with one being most favorable and five being least favorable.

We create the recommendation indicator ($dX_3(t)$) of the first 30 minutes after the recommendation update to see if it can serve as another predictor of jumps. Analyst recommendation changes are often published outside of usual market hours. In such cases, the indicator will respond during the first 30 minutes of trading on the next day if the release was after 4:00pm and before midnight or same day if the release was after midnight and before 9:30am. As shown in

Table 8, we again find that the p-values of the JPT statistics for all stocks are less than 5%, leading us to conclude that analyst recommendation information is another jump predictor.

4 Implications of Jump Predictor Identification and Irregular Jump Dynamics

Given the presence of two different types of risk in asset returns, namely jump and volatility risk, a thorough investigation, based on our new test, on what kind of information covariates affect the stochastic behavior of jump arrivals will help identify inter-temporal dynamics of higher moments of asset return distributions. Development of models with such characteristics implies, in general, the need for different portfolio and risk management strategies, as well as new pricing and hedging strategies. In particular, since derivative securities are affected more strongly than others by the higher moments of underlying asset returns due to jumps, stochastic jump arrivals imply a stronger impact on any financial management dealing with those securities. More specifically, uncertainty is reduced from increasing the predictability involved with discontinuous jump risks, leading to enhanced asset pricing models. This benefit is also emphasized in Bates (2003).

The majority of existing asset pricing models using both volatility and jump risks usually include stochastic volatility risk, but not many of them include stochastic jump risk. [see Bakshi, Cao, and Chen (1997), Bates (1996), Aït-Sahalia (2004), Eraker, Johannes, and Polson (2003).] Recognizing that the estimated parameter values are not stable over time under stochastic volatility but constant intensity jump arrivals [see Bates (1996)], more recent studies begin to allow jump arrival rates to depend on variables such as latent volatility, latent jump size, or market information in both an affine and a non-affine format. [see Pan (2002), Chernov, Gallant, Ghysels, and Tauchen (2003), Dubinsky and Johannes (2006), for example.] Because of the difficulty of the empirical identification of jump dynamics, incorporating relevant information in asset pricing models was not trivial, even if the nature of jumps in financial time series would demand more of such incorporation. Our JPT resolves this problem by providing a systematic way to set up the models with time-varying jump risks.

4.1 Simulation Study on the Impact of Jump Predictor Identification

We illustrate in this subsection economic implication of our empirical findings with a more specific example. In particular, we show how jump clustering has an impact on the evaluation of value at risk, denoted as "VaR" by simulation. VaR has been used extensively as a quantitative benchmark to disclose the financial risk of banks, securities firms, or any bundle of their portfolios in a liquid market environment. The definition of VaR is the loss over a given future time horizon s that is exceeded with a probability p for the given confidence level at $1-p$. Since it is the pth percentile of future return distribution at time s, it depends critically on higher moments, on which the presence of jumps would have stronger influence. The choice of s and p can be arbitrary. In practice, for s, a day, a week, or some other accounting period is a common choice, and p can be 1% or 5% depending on user's objectives. [see Duffie and Pan (1997).]

Under the jump clustering in normal trading we find in Subsection 3.3.2, the future intensity of jump at s is expected to be higher than under a static jump structure if a jump had occurred in previous hours. Therefore, one expects fatter tails in return distributions at s with the dynamic intensity, given that other functions are the same in the comparison. The fatter tail in the likelihood function makes VaR larger, because the pth percentile of the return distribution at s with dynamic intensity with jump clustering is lower.[12] It follows that the evidence implies higher VaR.

[12] p is usually less than 10.

In Table 9, we show examples with simulated VaR of an individual equity and portfolios of put options on that equity when jumps cluster. In this study, we compare two cases, one with a static and one with a dynamic jump intensity to show the effect. For series generation, we simulate 2,000 series from a jump diffusion model for equity returns. It is assumed that other functions are the same under both models. We set the spot volatility σ at 30%, standard deviation of jump sizes σ_y at 60%, static jump intensity λ at 10 jumps per year, and the coefficient θ_1 for jump clustering at 5%.[13] Current price is set at 100, and its future return is derived from the simulation. Strike prices of put options in the examples are set at 120 and 80 for in-the-money and out-of-the-money options, respectively. Maturity is set at 1 year. The values for VaR in this exercise are the empirical percentiles of the simulated return distributions. Table 9 illustrates various cases of VaR and its put options at given conditions. Table 9 shows that jump clustering consistently yields higher values of VaR for all cases.

5 Concluding Remarks

We investigate dynamics of stochastic jump risks in asset pricing model. To study its stochastic features in relation to time-varying financial market information flow, we first introduce a new empirical test to identify relevant information covariates that significantly affect irregular jump arrivals. Using this test, we investigate the predictability of higher moments of asset return distributions. The test takes a two-stage semi-parametric approach of applying the nonparametric jump detection test to first distinguish jump arrivals in Stage I and the partial maximum likelihood test in Stage II to identify the significant predictors. Monte Carlo simulation proves that our test can distinguish the effect of informational predictors for jump arrivals using high frequency

[13] We choose a small value for the coefficient with which we still prove the effect is strong.

observations. As long as high frequency observations are available for target returns and proposed jump predictors during a sufficiently long time period, this test can be applied to any type of financial time series. Hence, this methodological contribution applies to all kinds of markets, including equity, bond, foreign currency, and their corresponding options and other derivatives markets.

Using the new test, we perform an empirical analysis on major U.S. individual stocks that are actively traded on the NYSE. These results show that the earnings announcement effect on jump arrivals is stronger in the morning when information is accumulated and reflected in prices than during the remainder of the day. We also show evidence of jump clustering during normal trading hours in equity markets — more jumps during preceding trading hours lead to an increased intensity of future jump arrivals. Interestingly, we find more jumps are triggered by negative jumps than positive jumps. We also report evidence that publication of analyst recommendations is another jump predictor for individual stocks. These findings have several economic implications for portfolio and risk management as well as asset pricing. In particular, we show how strongly this evidence will have an impact on the VaR of equity returns and their corresponding options, among other issues.

Our test is designed in this paper to detect and find the information covariate for only non-homogeneous Poisson-type rare jumps in jump diffusion frameworks. With a minor modification in estimating the instantaneous volatility in Stage I, a similar method is applicable to discrete observations from jump diffusion models with infinite-activity Lévy jump processes. [see Lee and Hannig (2007) for more detailed explanation.] Especially, if the purpose of the study is to understand the dynamics of specific types of jumps (for instance, large-sized jumps) from infinite-activity jump processes, the proposed test in Stage II we set forth in this paper is also applicable.

Another extension of our empirical studies can be for other markets, such as bond and foreign exchange markets and to find cross-sectional relations between several different markets.

Appendix

A.1. Assumption C on $\mu(t)$ and $\sigma(t)$ in equation (1)

The following assumptions are imposed on the drift $\mu(t)$ and spot volatility $\sigma(t)$ in the stochastic differential equation (equation (1)) for the asset price $S(t)$ [see Lee and Mykland (2006)].

For any $\epsilon > 0$ and $\Delta t = t_{i+1} - t_i$,

$$\textbf{C.1} \ \sup_i \sup_{t_i \leq u \leq t_{i+1}} |\mu(u) - \mu(t_i)| = O_p(\Delta t^{\frac{1}{2}-\epsilon}) \tag{14}$$

$$\textbf{C.2} \ \sup_i \sup_{t_i \leq u \leq t_{i+1}} |\log \sigma(u) - \log \sigma(t_i)| = O_p(\Delta t^{\frac{1}{2}-\epsilon}) \tag{15}$$

A.2. Assumption D on $\Lambda_\theta(t)$ in equation (2)

The following assumptions are imposed on $\Lambda_\theta(t)$, which is a modified version of Condition VI.1.1. of Andersen, Borgan, Gill, and Keiding (1992). Denote by θ_0 the true value of parameter and θ the free parameter. Let T be a given terminal time, $0 < T \leq \infty$ and n be the number of observations within the terminal time T.

D1) There exists a neighborhood Θ_0 of θ_0 such that for all n and $\theta \in \Theta_0$, $\log d\Lambda_\theta(t)$ and $d\Lambda_\theta(t)$ are three times differentiable with respect to $\theta \in \Theta_0$.

D2) There exist finite functions $\sigma_{jl}(\theta)$ defined on Θ_0 such that for all j, l,

$$\frac{1}{n}\int_0^\tau \{\frac{\partial}{\partial\theta_j} \log d\Lambda_{\theta_0}(t)\}\{\frac{\partial}{\partial\theta_l} \log d\Lambda_{\theta_0}(t)\} d\Lambda_{\theta_0}(t)dt \xrightarrow{p} \sigma_{jl}(\theta_0)$$

as $n \to \infty$. Moreover, the matrix $\Sigma = \{\sigma_{jl}(\theta_0)\}$ is positive definite.

D3) For all j and $\epsilon > 0$, we have

$$\frac{1}{n}\int_0^\tau \{\frac{\partial}{\partial\theta_j} \log d\Lambda_{\theta_0}(s)\}^2 I\left(|\frac{1}{\sqrt{n}}\frac{\partial}{\partial\theta_j} \log d\Lambda_{\theta_0}(s)| > \epsilon\right) d\Lambda_{\theta_0}(s)ds \xrightarrow{p} 0$$

as $n \to \infty$.

D4) For any n, there exist G_n and H_n such that

$$\sup_{\theta \in \Theta_0} |\frac{\partial^3}{\partial \theta_j \partial \theta_l \partial \theta_m} d\Lambda_\theta(t)| \leq G_n(t)$$

and

$$\sup_{\theta \in \Theta_0} |\frac{\partial^3}{\partial \theta_j \partial \theta_l \partial \theta_m} \log d\Lambda_\theta(t)| \leq H_n(t)$$

for all j, l, m. And

$$\frac{1}{n}\int_0^\tau G_n(t)dt,\ \frac{1}{n}\int_0^\tau H_n(t)d\Lambda_{\theta_0}(t)dt,\ \frac{1}{n}\int_0^\tau \{\frac{\partial^2}{\partial \theta_j \partial \theta_l} \log d\Lambda_{\theta_0}(t)\}^2 d\Lambda_{\theta_0}(t)dt$$

all converge in probability to finite quantities as $n \to \infty$, and for all $\epsilon > 0$,

$$\frac{1}{n}\int_0^\tau H_n(t) I\left(\sqrt{\frac{H_n(t)}{n}} > \epsilon\right) d\Lambda_{\theta_0}(t)dt \xrightarrow{p} 0$$

A.3. Proof of Proposition 1

We go back to the categorization of all time intervals into two different types as discussed earlier in subsection 2.1.1. The first set of times includes times $i \in \{1, .., n\}$ with actual jumps in the interval (t_{i-1}, t_i), which we call JT_n. The other set of times includes times $i \in \{1, .., n\}$ without actual jumps in the interval (t_{i-1}, t_i), which we call NJT_n. The two sets are mutually exclusive so that $JT_n \bigcup NJT_n = \{1, 2, .., n-1, n\}$ and $JT_n \bigcap NJT_n = \emptyset$. From equation (7), we can further decompose the approximate likelihood function into two different mutually exclusive parts for actual jump times and non-jump times as follows:

$$L(\theta|\mathcal{F}_T) = \underbrace{\prod_{1 \leq i \leq n, i \in JT_n} d\Lambda_\theta(t_i)^{dJ(t_i)}}_{(16.1)} \underbrace{\prod_{1 \leq i \leq n, i \in JT_n} (1 - d\Lambda_\theta(t_i))^{1-dJ(t_i)}}_{(16.2)}$$

$$\times \underbrace{\prod_{1 \le i \le n, i \in NJT_n} d\Lambda_\theta(t_i)^{dJ(t_i)}}_{(16.3)} \underbrace{\prod_{1 \le i \le n, i \in NJT_n} (1 - d\Lambda_\theta(t_i))^{1-dJ(t_i)}}_{(16.4)} \tag{16}$$

$$\text{where } \Lambda_\theta(t) = \gamma(t, X(t); \theta).$$

The second (16.2) and third (16.3) products are 1 under the full observation of jumps. Hence, it is enough to show that both of these two products, (16.2) and (16.3), based on partial observations, become 1, with probability 1, as $\Delta t \to 0$, so that the other two products based on partial observations match the corresponding ones, (16.1) and (16.4). From Theorems 3 and 4 of Lee and Mykland (2006), we know that when N is the number of jumps during the time horizon,

$$P\left(\text{for all } i \in JT_n,\, d\hat{J}(t_i) = 1\right) \approx 1 - \frac{2}{\sqrt{2\pi}} y_n N + o(y_n^2 N) \to 1 \text{ as } \Delta t \to 0, \tag{17}$$

where y_n is defined in equation (14) in Lee and Mykland (2006). Therefore,

$$P\left(\prod_{1 \le i \le n, i \in JT_n} \left(1 - d\hat{\Lambda}_\theta(t_i)\right)^{1-d\hat{J}(t_i)} = 1\right) \approx 1 - \frac{2}{\sqrt{2\pi}} y_n N + o(y_n^2 N) \to 1 \text{ as } \Delta t \to 0. \tag{18}$$

Similarly,

$$P\left(\text{for all } i \in NJT_n,\, d\hat{J}(t_i) = 0\right) = 1 - P\left(\text{for any } i \in NJT_n,\, d\hat{J}(t_i) = 1\right)$$

$$= 1 - exp(\beta_n) + o(\beta_n) \to 1 \text{ as } \Delta t \to 0, \tag{19}$$

where β_n is defined as in equation (40) in Lee and Mykland (2006). Hence,

$$P\left(\prod_{1 \le i \le n, i \in NJT_n} d\hat{\Lambda}_\theta(t_i)^{d\hat{J}(t_i)} = 1\right) = 1 - exp(\beta_n) + o(\beta_n) \to 1 \text{ as } \Delta t \to 0. \tag{20}$$

Therefore, equation (9) holds, because

$$P\left(L(\theta|\mathcal{F}_T) \neq PL(\theta|\mathcal{F}_T)|N\right) = 1 - P\left(L(\theta|\mathcal{F}_T) = PL(\theta|\mathcal{F}_T)|N\right)$$

$$= 1 - P\left(\prod_{1 \le i \le n, i \in NJT_n} d\hat{\Lambda}_\theta(t_i)^{d\hat{J}(t_i)} = 1\right) \times P\left(\prod_{1 \le i \le n, i \in JT_n} \left(1 - d\hat{\Lambda}_\theta(t_i)\right)^{1-d\hat{J}(t_i)} = 1\right)$$

$$\approx 1 - \left(1 - \frac{2}{\sqrt{2\pi}} y_n + o(y_n^2)\right) \times \left(1 - exp(\beta_n) + o(\beta_n)\right) \to 0 \text{ as } \Delta t \to 0, \tag{21}$$

and by the definition of the product-integration, the second statement consequently holds.

A.4. Proof of Proposition 2

Given the assumption C, we know that as $\Delta t \to 0$, for any θ, $L(\theta) - PL(\theta) \to 0$ in probability from the first part of Proposition 1. Here, let $\mathcal{U}_L(\theta)$ and $\mathcal{U}_{PL}(\theta)$ be the score functions based on $L(\theta)$ and $PL(\theta)$. Then, the two estimators, $\hat{\theta}_{L,n}$ and $\hat{\theta}_{PL,n}$ such that $\mathcal{U}_L(\hat{\theta}_{L,n}) = 0$ and $\mathcal{U}_{PL}(\hat{\theta}_{PL,n}) = 0$ are asymptotically equivalent. In other words, as $\Delta t \to 0$ (as $n \to 0$), $\hat{\theta}_{L,n} - \hat{\theta}_{PL,n} \to 0$ in probability: this is proved by contradiction. Now, according to the Slutsky Theorem as in Ferguson (1996), it is enough to show that the estimator based on L, $\hat{\theta}_{L,n}$, is consistent and converges in law to a normal distribution around its mean θ_0. For this part, we apply a modified version of proofs for Theorem VI.1.1. and VI.1.2 in Andersen, Borgan, Gill, and Keiding (1992). A sketch of the proofs are as follows. Due to a Taylor expansion, $1 - d\Lambda_\theta(t) = exp(-d\Lambda_\theta(t))$, $\mathcal{U}_L(\theta)$ can be written as

$$\mathcal{U}_L(\theta) = \int_0^{\cdot} \frac{\partial}{\partial \theta} \log d\Lambda_\theta(s) dM(s),$$

where $M(t) = J(t) - \int_0^t d\Lambda_\theta(s)ds$ and is a local square integrable martingale. Here, we first apply Lenglart's inequality to establish the existence of a consistent estimator that is the solution for the score function. Second, we use the Martingale Central Limit Theorem to establish the convergence of estimators in distribution to normal. Lastly, it is obvious that our JPT, a version of the Wald test, asymptotically follows χ^2 distribution with q degrees of freedom.

Alternative to the proof given above is to consider two equivalent probability measures $\mathcal{P}$ and $\mathcal{P}_{PL}$. $\mathcal{P}$ is the true (latent) data-generating measure for $\widetilde{L(\theta)}$ in continuous time (same as in page 5) and $\mathcal{P}_{PL}$ is the observable data-generating measure for $PL(\theta)$ in discrete time. Instead of going through $L(\theta)$, the above weak convergence proof can be directly applied on $\mathcal{U}_{PL}(\theta)$, because of the convergence of $\mathcal{P}_{PL}$ to $\mathcal{P}$.

References

Aït-Sahalia, Y., 2002, "Telling from Discrete Data Whether the Underlying Continuous-Time Model is a Diffusion," *Journal of Finance*, 57, 2075–2112.

Aït-Sahalia, Y., 2004, "Disentangling Diffusion from Jumps," *Journal of Financial Economics*, 74, 487–528.

Aït-Sahalia, Y., and J. Jacod, 2006a, "Testing for Jumps in a Discretely observed Process," Working Paper, Department of Economics, Princeton Usniversity.

Aït-Sahalia, Y., and J. Jacod, 2006b, "Volatility Estimators for Discretely Sampled Lévy processes," *Annals of Statistics, forthcoming.*

Aït-Sahalia, Y., P. A. Mykland, and L. Zhang, 2006, "Ultra High Frequency Volatility Estimation with Dependent Microstructure Noise," Working Paper, Department of Economics, Princeton University.

Amihud, Y., and H. Mendelson, 1987, "Trading Mechanisms and Stock Returns: An Empirical Investigation," *Journal of Finance*, 42, 533–553, Papers and Proceedings of the Forty-Fifth Annual Meeting of the American Finance Association, December 1986.

Andersen, P. K., O. Borgan, R. D. Gill, and N. Keiding, 1992, *Statistical Models based on Counting Processes*, Springer-Verlag, New York Berlin Heidelberg.

Andersen, T., T. Bollerslev, F. Diebold, and H. Ebens, 2001, "The Distribution of Realized Stock Return Volatility," *Journal of Financial Economics*, 61, 43–76.

Andersen, T., T. Bollerslev, F. Diebold, and C. Vega, 2003, "Micro Effects Macro Announcements: Real-Time Price Discovery in Foreign Exchange," *American Economic Review*, 93, 38–62.

Andersen, T. G., 1996, "Return Volatility and Trading Volume: An Information Flow Interpretation of Stochastic Volatility," *Journal of Finance*, 51, 169–204.

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

Bakshi, G., C. Cao, and Z. Chen, 1997, "Empirical Performance of Alternative Option Pricing Models," *Journal of Finance*, 52, 2003–2049.

Barndorff-Nielsen, O. E., and N. Shephard, 2004, "Power and Bipower Variation with Stochastic Volatility and Jumps," *Journal of Financial Econometrics*, 2, 1–48.

Barndorff-Nielsen, O. E., and N. Shephard, 2005, "Variation, jumps, market frictions and high frequency data in financial econometrics," Working Paper, University of Oxford.

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

Barndorff-Nielsen, O. E., N. Shephard, and M. Winkel, 2005, "Limit Theorems for Multipower Variation in the Presence of Jumps," Working Paper, Oxford University.

Bates, D. S., 1996, "Jumps and Stochastic Volatility: Exchange Rate Processes Implicit in Deutsch Mark Options," *Reivew of Financial Studies*, 9, 69–107.

Bates, D. S., 2003, "Empirical Option Pricing: A Retrospection," *Journal of Econometrics*, 116, 387–404.

Bollerslev, T., R. Y. Chou, and K. F. Kroner, 1992, "ARCH Modeling in Finance," *Journal of Econometrics*, 52, 5–59.

Campbell, J. Y., A. W. Lo, and A. C. MacKinlay, 1997, *The Econometrics of Financial Markets*, Princeton University Press, Princeton, NJ.

Carr, P., and L. Wu, 2006, "Stochastic Skew in Currency Options," *Journal of Financial Economics, forthcoming.*

Chan, K., P. Y. Chung, and W.-M. Fong, 2002, "The information role of stock and option volume," *Review of Financial Studies*, 15, 1949–1975.

Chernov, M., A. R. Gallant, E. Ghysels, and G. Tauchen, 2003, "Alternative Models for Stock Price Dynamics," *Journal of Econometrics*, 116, 225–257.

Dubinsky, A., and M. Johannes, 2006, "Earnings announcements and Option Prices," Working Paper, Graduate School of Business, Columbia University.

Duffie, D., and J. Pan, 1997, "An Overview of Value at Risk," *Journal of Derivatives*, pp. 7–49, reprinted in Options Markets, edited by G. Constantinides and A. G. Malliaris, London: Edward Elgar, 2001.

Easley, D., M. O'Hara, and P. Srinivas, 1998, "Option Volume and Stock Prices: Evidence on Where Informed Traders Trade," *Journal of Finance*, 53, 431–465.

Eraker, B., M. Johannes, and N. Polson, 2003, "The Impact of Jumps in Volatility and Returns," *Journal of Finance*, 53, 1269–1300.

Fama, E., and K. French, 1988, "Dividend yields and expected stock returns," *Journal of Financial Economics*, 22, 3–26.

Fama, E., and K. French, 1990, "Business conditions and expected returns on stocks and bonds," *Journal of Financial Economics*, 25, 23–49.

Ferguson, T. S., 1996, *A Course in Large Sample Theory*, Chapman and Hall, London.

Galambos, J., 1978, *The Asymptotic Theory of Extreme Order Statistics*, John Wiley and Sons, Inc.

Green, T. C., 2007, "The Value of Client Access to Analyst Recommendations," *Journal of Financial and Quantitative Analysis, forthcoming.*

Heston, S., 1993, "A Closed-Form Solution for Options with Stochastic Volatility with Applications to Bonds and Currency Options," *Review of Financial Studies*, 6, 327–343.

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

Hull, J., and A. White, 1987, "The Pricing of Options on Assets with Stochastic Volatilities," *Journal of Finance*, 42, 281–300.

Jiang, G. J., and R. Oomen, 2005, "A New Test for Jumps in Asset Prices," Working Paper, Eller College of Management, University of Arizona.

Johannes, M. S., R. Kumar, and N. Polson, 1999, "State Dependent Jump Models: How do US Equity Indices Jump?," Working Paper, Graduate School of Business, Univeristy of Chicago.

Keim, D. B., and R. F. Stambaugh, 1986, "Predicting Returns in the Stock and Bond Markets," *Journal of Financial Economics*, 17, 357–390.

Kloeden, P. E., and E. Platen, 1992, *Numerical Solution of Stochastic Differential Equations*, Springer-

Verlag.

Lee, S., 2007, "Financial Risks Decompositions," Working Paper, Georgia Institute of Technology.

Lee, S., and J. Hannig, 2007, "Detecting Jumps from Lévy Jump Diffusion," Working Paper, Georgia Institute of Technology.

Lee, S., and P. A. Mykland, 2006, "Jumps in Financial Markets: A New Nonparametric Test and Jump Dynamics," *Review of Financial Studies, forthcoming.*

Lo, A. W., and J. Wang, 1995, "Implementing Option Pricing Models When Asset Returns Are Predictable," *Journal of Finance*, 50, 87–129.

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

Merton, R. C., 1992, *Continuous-Time Finance*, Blackwell.

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

Protter, P., 2004, *Stochastic Integration and Differential Equations: A New Approach*, Spinger, New York.

Stein, E. M., and J. C. Stein, 1991, "Stock Price Distributions with Stochastic Volatility: An Analytic Approach," *Review of Financial Studies*, 4, 727–752.

Tookes, H. E., 2006, "Information, Trading and Product Market Interactions: Cross-sectional Implications of Informed Trading," *Journal of Finance, forthcoming.*

Womack, K. L., 1996, "Do brokerage analysis' recommendations have investment value?," *Journal of Finance*, 51, 137–167.

Table 1: **A Result from the Jump Predictor Test (JPT)** †

$\hat{\theta}_0$	Standard Error($\hat{\theta}_0$)	$\hat{\theta}_1$	Standard Error ($\hat{\theta}_1$)
0.0101	6.3996e-004	0.9066	0.0798
95% Confidence Interval for θ_0		95% Confidence Interval for θ_1	
[0.0088,0.0113]		[0.7502,1.0630]	
JPT Statistic on $X(t)$		P-value	
129.1164		0	

† This table contains an example of simulated results from the semi-parametric JPT and its related tests. The purpose of this predictor test is to learn whether a chosen information covariate affects the stochastic jump intensity. The encompassing model is $d\log S(t) = \mu(t)dt+\sigma(t)dW(t)+Y(t)dJ(t)$ with non-homogeneous Poisson process $J(t)$. Constant Volatility $\sigma(t) = \sigma$ is set at 30%. Jump sizes are set in comparison with volatility level. The standard deviation σ_y of jump sizes is set at three times σ. We suppose $dX(t)$ to be an indicator function of monthly information release, which is related to jump arrival in an affine function: $d\Lambda_\theta(t) = \theta_0 + \theta_1 dX(t)$. θ_0 and θ_1 are assumed to be 0.01 and 0.99 so that the expected jump intensity per year is around 5%. The significance level α equal to 5% and 15 minute return data are used in Stage I.

Table 2: **Averaged Simulation Results from the Jump Predictor Test (JPT)**[†]

$\sigma_y(t) = 3 \times \sigma(t)$			
$\hat{\theta}_0$	Standard Error($\hat{\theta}_0$)	$\hat{\theta}_1$	Standard Error ($\hat{\theta}_1$)
0.0098	6.3140e-004	0.8934	0.0841
JPT Statistic on $X(t)$		P-value	
118.5533		4.9066e-011	
$\sigma_y(t) = 2 \times \sigma(t)$			
$\hat{\theta}_0$	Standard Error($\hat{\theta}_0$)	$\hat{\theta}_1$	Standard Error ($\hat{\theta}_1$)
0.0096	6.2581e-004	0.8782	0.0892
JPT Statistic on $X(t)$		P-value	
105.7017		6.4140e-011	
$\sigma_y(t) = 1 \times \sigma(t)$			
$\hat{\theta}_0$	Standard Error($\hat{\theta}_0$)	$\hat{\theta}_1$	Standard Error ($\hat{\theta}_1$)
0.0096	6.2359e-004	0.8813	0.0875
JPT Statistic on $X(t)$		P-value	
111.2553		1.4296e-008	

[†] This table contains averaged simulated results from the semi-parametric JPT shown in Table 1 for different jump size distributions. All the figures in this table are averaged ones across different simulation runs. The purpose of this table is to show that jump size changes do not affect the results of tests to determine the effect of information covariates. The encompassing model is $d \log S(t) = \mu(t)dt + \sigma(t)dW(t) + Y(t)dJ(t)$. Constant volatility sets $\sigma(t) = \sigma$ at 30%. Jump sizes are set in comparison with volatility level. The standard deviation σ_y of jump size distribution is set at three different levels: $3 \times \sigma$, $2 \times \sigma$, and $1 \times \sigma$. We suppose $dX(t)$ to be an indicator of monthly information release and is related to jump arrival in an affine function: $d\Lambda_\theta(t) = \theta_0 + \theta_1 dX(t)$. θ_0 and θ_1 are assumed to be 0.01 and 0.99 so that the expected jump intensity per year is around 5%. The significance level α equals to 5% and 15-minute return data are used in Stage I.

Table 3: **Jump Size (in %) Distributions of DJIA Individual Equities** †

Ticker	ReJ	Mean	Med	Std	Skew	Kurt	Min	Max	Lq	Uq	MinAb
AA	105	-0.079	1.269	3.368	-0.257	3.429	-9.901	10.436	-2.239	2.233	0.916
AIG	97	-0.131	-0.789	3.031	-0.501	4.542	-10.852	7.042	-1.893	1.787	0.705
AXP	110	-0.036	0.788	3.016	-2.705	17.628	-19.034	5.640	-1.459	1.712	0.529
BA	122	0.087	1.222	2.921	-2.829	20.192	-20.177	5.928	-1.801	1.897	0.832
C	117	-0.162	0.571	2.721	-0.358	6.549	-10.510	11.227	-1.381	1.386	0.570
CAT	94	-0.690	1.129	7.533	-7.877	71.167	-68.434	5.640	-1.858	2.002	0.790
DD	107	0.067	0.928	2.076	-0.188	2.289	-5.515	4.628	-1.496	1.629	0.807
DIS	109	-0.422	-1.032	3.973	-1.386	11.496	-21.426	14.299	-1.961	1.912	0.692
GE	113	0.381	1.008	2.559	-0.629	4.823	-10.423	7.388	-1.534	2.114	0.673
GM	135	0.564	1.382	3.255	-0.178	5.704	-11.838	11.939	-1.855	2.323	0.942
HD	108	-0.349	-1.088	3.413	-0.748	5.702	-12.405	10.159	-1.970	1.924	0.827
HON	106	0.255	1.384	3.766	-1.929	8.716	-16.089	6.068	-1.852	2.546	0.917
HPQ	118	-0.420	-1.323	4.002	-0.264	5.524	-16.962	9.738	-2.289	1.980	0.672
IBM	117	0.001	-0.581	2.820	0.306	7.258	-11.251	12.300	-1.553	1.339	0.510
JNJ	127	-0.163	-0.593	2.328	-3.372	27.980	-17.931	5.550	-1.299	1.275	0.411
JPM	121	-0.187	-0.714	3.233	-0.392	4.372	-12.085	8.196	-1.846	1.761	0.506
KO	79	-0.463	-0.932	2.393	-0.847	4.464	-8.455	4.015	-1.559	1.308	0.582
MCD	125	-0.223	-0.809	2.686	-0.904	5.883	-13.124	5.574	-2.062	1.776	0.762
MMM	65	-1.029	0.848	8.888	-7.207	56.229	-69.600	5.165	-1.738	1.515	0.808
MO	138	-0.351	-1.009	2.975	-0.312	5.926	-13.039	10.610	-1.883	1.530	0.627
MRK	138	-0.467	-0.933	3.598	-4.489	39.389	-31.127	6.556	-1.832	1.557	0.864
PFE	147	-0.142	0.708	3.231	-1.690	12.348	-18.786	10.899	-1.809	1.706	0.708
PG	89	-0.782	0.629	7.524	-8.635	79.228	-69.333	3.271	-1.240	1.273	0.629
T	135	0.521	-0.928	6.888	6.384	60.567	-16.893	65.467	-2.328	2.379	0.589
VZ	102	-0.393	-0.959	2.196	-0.381	2.299	-6.392	3.533	-1.878	1.465	0.676
UTX	97	-0.539	0.889	7.593	-7.978	72.682	-69.697	6.677	-1.672	1.606	0.747
WMT	97	0.156	0.936	1.909	-0.246	2.407	-5.030	4.089	-1.371	1.569	0.759
XOM	90	-1.264	-1.101	7.215	-8.411	76.938	-66.712	3.496	-1.842	1.272	0.725
Average	111	-0.223	0.032	3.968	-2.072	22.347	-23.679	9.483	-1.768	1.742	0.706
S.E.	3.603	0.08	0.187	0.3827	0.6325	5.036	4.172	2.152	0.053	0.067	0.025

† The table shows summary statistics that characterize the distributions of jump size [$Y(t)$ in equation (1)] of individual stocks for the DJIA index. Reported are those traded on the NYSE. We exclude Intel (INTC) and Microsoft (MSFT) since the trading mechanism of the NASDAQ stock exchange makes their intraday observation patterns different, such that their comparison to the rest of the sample is not appropriate. The sample under consideration extends 5 years from July, 2001 to June, 2006 for a total of 1256 trading days. We use 15-minute returns based on transaction data to detect jump arrivals in Stage I. Ticker denotes the ticker name of each company. ReJ denotes the number of jumps realized (detected) during our sample period of 5 years. Mean, Med, Std, Skew, Kurt, Min, Max, Lq, Uq, and MinAbs denote, respectively, the mean, median, standard deviation, skewness, kurtosis, lower quartile, upper quartile, and minimum of absolute value of observed jump sizes. Averages and S.E. are for the cross-sectional averages of the corresponding columns. The significance level $\alpha = 5\%$ is used in Stage I.

Table 4: **Jump Counting: Average Number of Detected Jumps (ANDJ)**[†]

Ticker	per year	per month	per day	Ticker	per year	per month	per day
AA	21.0	1.75	0.083	AIG	19.4	1.62	0.077
AXP	22.0	1.83	0.087	BA	24.4	2.03	0.096
C	23.4	1.95	0.092	CAT	18.8	1.57	0.074
DD	21.4	1.78	0.084	DIS	21.8	1.82	0.086
GE	22.6	1.88	0.089	GM	27.0	2.25	0.107
HD	21.6	1.80	0.085	HON	21.2	1.77	0.084
HPQ	23.6	1.97	0.093	IBM	23.4	1.95	0.092
JNJ	25.4	2.12	0.100	JPM	24.2	2.02	0.096
KO	15.8	1.32	0.062	MCD	25.0	2.08	0.099
MMM	13.0	1.08	0.051	MO	27.6	2.30	0.109
MRK	27.6	2.30	0.109	PFE	29.4	2.45	0.116
PG	17.8	1.48	0.070	T	27.0	2.25	0.107
VZ	20.4	1.70	0.081	UTX	19.4	1.62	0.077
WMT	19.4	1.62	0.077	XOM	18.0	1.50	0.071

per year	Average	22.200	S.E.	3.814
per month	Average	1.850	S.E.	0.318
per day	Average	0.0876	S.E.	0.015

[†] This table includes the yearly, monthly, and daily average number of detected jumps for each individual component stock of the DJIA, and their cross-sectional averages. Reported are those stocks traded on the NYSE for 5 years from July 1, 2001 to June 30, 2006, for a total of 1256 trading days. We exclude Intel (INTC) and Microsoft (MSFT) since the trading mechanism of the NASDAQ stock exchange makes their intraday observation patterns different such that comparison to the rest of the sample is not appropriate. The 15-minute returns from transaction prices are used to count the number of detected jumps. The significance level α in Stage I is 5%.

Table 5: **At What Time do Jumps Occur more often?**†

Ticker	9:30am	10:00am	10:30am	11:30am	12:30pm	1:30pm	2:30pm	3:30pm	4:00pm
AA	15.23	78.09	4.76	0.95	0	0	0	0	0.95
AIG	26.80	51.54	5.15	6.18	5.15	2.06	2.06	0	1.03
AXP	20.90	52.72	11.81	2.72	1.81	0.90	2.72	4.54	1.81
BA	27.04	57.37	11.47	2.45	0	0	0	0.81	0.81
C	34.18	47.86	8.54	0.85	1.70	0	1.70	2.56	2.56
CAT	15.95	68.08	7.44	5.31	1.06	0	0	1.06	1.06
DD	19.62	66.35	5.60	2.80	3.73	0	0	0.93	0.93
DIS	28.44	60.55	4.58	0.91	0.91	0	0	0.91	3.66
GE	54.86	31.85	7.07	0.88	1.76	0	0	0.88	2.65
GM	18.51	59.25	10.37	2.96	0	1.48	3.70	2.22	1.48
HD	15.74	66.66	6.48	3.70	0	0.92	2.77	3.70	0
HON	18.86	65.09	6.60	1.88	1.88	0	3.77	0	1.88
HPQ	29.66	61.01	2.54	2.54	2.54	0.84	0	0	0.84
IBM	66.66	28.20	0	0.85	0.85	0	0	0.85	2.56
JNJ	26.77	60.62	5.51	3.14	0	0.78	0.78	0.78	1.57
JPM	43.80	39.66	4.95	4.13	0	0.82	2.47	3.30	0.82
KO	35.44	54.43	1.26	0	1.26	2.53	1.26	3.79	0
MCD	29.60	52.80	2.40	5.60	1.60	1.60	4.00	1.60	0.80
MMM	10.76	73.84	7.69	3.07	1.53	1.53	1.53	0	0
MO	15.21	52.89	10.86	6.52	2.89	2.17	5.79	1.44	1.44
MRK	8.69	72.46	3.62	2.89	3.62	2.17	0.72	4.34	1.44
PFE	25.17	63.26	4.08	1.36	2.04	1.36	0	1.36	0.68
PG	34.83	56.17	5.61	0	1.12	0	0	0	2.24
T	11.11	74.07	6.66	1.48	2.96	0	0.74	1.48	1.48
VZ	27.45	63.72	3.92	0	1.96	0	2.94	0	0
UTX	20.61	65.97	6.18	5.15	0	0	2.06	0	0
WMT	13.40	78.35	4.12	0	1.03	0	1.03	1.03	1.03
XOM	33.33	56.66	3.33	2.22	1.11	1.11	1.11	0	1.11
Mean	26.02	59.26	5.80	2.52	1.52	0.72	1.47	1.34	1.24

† The table reports the percentages of detected individual equity jumps occurred in specific time intervals of a trading day among all occurrences during our sample period from July 1, 2001 to June 30, 2006 for a total of 1256 trading days. We divided the NYSE trading day (9:30am to 4:00pm) into nine categories. Column 9:30am is for the market opening, and columns 10:00am, 10:30am, 11:30am, 12:30pm, 1:30pm, 2:30pm, and 3:30pm are ending times from the previous columns. For instance, 1:30pm contains the jumps that arrived from 12:30pm till 1:30pm.

Table 6: **Effect of Earnings Announcement Information Release on Individual Component Stocks in the DJIA Index** †

Ticker: C (Citigroup Inc)			
Covariate ($X_0(t)$)	Opening on EADs	Morning on EADs	EADs
$\hat{\theta}_0$	0.0021	0.0021	0.0021
Standard Error($\hat{\theta}_0$)	2.0807e-4	2.0641e-4	2.0996e-4
$\hat{\theta}_1$	0.1479	0.0854	0.0108
Standard Error($\hat{\theta}_1$)	0.0399	0.0223	0.0034
JPT Statistic on $X(t)$	13.7187	14.6159	9.8925
P-value	2.1232e-4	1.3180e-4	0.0017
Ticker: DIS (Walt Disney Company)			
Covariate ($X_0(t)$)	Opening on EADs	Morning on EADs	EADs
$\hat{\theta}_0$	0.0019	0.0019	0.0019
Standard Error($\hat{\theta}_0$)	1.9584e-4	1.9404e-4	1.9642e-4
$\hat{\theta}_1$	0.1981	0.1106	0.0157
Standard Error($\hat{\theta}_1$)	0.0447	0.0250	0.0040
JPT Statistic on $X(t)$	19.6237	19.6191	15.4523
P-value	9.4292e-6	9.4518-6	8.4613e-5
Ticker: IBM (International Business Machine)			
Covariate ($X_0(t)$)	Opening on EADs	Morning on EADs	EADs
$\hat{\theta}_0$	0.0020	0.0020	0.0020
Standard Error($\hat{\theta}_0$)	2.0205e-4	2.0135e-4	2.0493e-4
$\hat{\theta}_1$	0.2230	0.1168	0.0156
Standard Error($\hat{\theta}_1$)	0.0467	0.0256	0.0040
JPT Statistic on $X(t)$	22.8122	20.8418	15.1278
P-value	1.7863e-6	4.9882e-6	1.0047e-4

† This table contains the empirical results of the semi-parametric JPT applied to three individual component stocks (C, DIS, and IBM) to determine whether earnings announcement information is a jump predictor. The sample period is 5 years from July 1, 2005 to June 30, 2006 and 15-minute returns are calculated from transaction prices. The significance level α used to detect jump arrivals in Stage I is equal to 5%. The model is $d\log S(t) = \mu(t)dt + \sigma(t)dW(t) + Y(t)dJ(t)$ with $J(t)$, a non-homogeneous Poisson process, whose stochastic intensity function is an affine function of $dX(t)$: $d\Lambda_\theta(t) = \theta_0 + \theta_1 dX_0(t)$. We set in this table $dX_0(t)$ to be indicators of opening times, morning times, and all trading times on earnings announcement dates and the next days for after trading hour announcements.

Table 7: **Jump Clustering in Individual Component Stocks of the DJIA Index**†

Ticker	JPT with $X_1(t)$	P-value	JPT with $X_2(t)$	P-value
AA	1.9931	0.1580	11.9811	5.3742e-004
AIG	11.9828	5.3695e-004	12.9815	3.1458e-004
AXP	6.9878	0.0082	10.9816	9.2021e-004
BA	9.9811	0.0016	13.9872	1.8406e-004
C	5.9888	0.0144	9.9819	0.0016
CAT	5.9911	0.0144	12.9822	3.1447e-004
DD	4.9914	0.0255	9.9836	0.0016
DIS	0.0144	0.0255	9.9833	0.0016
GE	2.9889	0.0838	7.9858	0.0047
GM	12.9866	3.1373e-004	21.9790	2.7565e-006
HD	4.9913	0.0255	16.9869	3.7638e-005
HON	10.9823	9.1984e-004	15.9878	6.3751e-005
HPQ	8.9834	0.0027	13.9877	1.8401e-004
IBM	4.9822	0.0256	10.9803	9.2085e-004
JNJ	6.9858	0.0082	11.9767	5.3870e-004
JPM	8.9830	0.0027	17.9843	2.2273e-005
KO	1.9948	0.1578	4.9937	0.0254
MCD	16.9845	3.7687e-005	21.9808	2.7539e-006
MMM	3.9900	0.0458	4.9877	0.0255
MO	25.9754	3.4580e-007	37.9679	7.1920e-010
MRK	14.9844	1.0840e-004	19.9801	7.8254e-006
PFE	9.9769	0.0016	19.9785	7.8317e-006
PG	3.9885	0.0458	6.9903	0.0082
T	15.9839	6.3883e-005	22.9782	1.6385e-006
VZ	4.9918	0.0255	9.9845	0.0016
UTX	9.9853	0.0016	12.9815	3.1458e-004
WMT	6.9894	0.0082	7.9880	0.0047
XOM	2.9912	0.0837	8.9877	0.0027

† This table illustrates that the more jumps that occur in previous trading hours, the stronger the impact on the likelihood of future jump arrivals. $dX_1(t)$ and $dX_2(t)$ are indicators of the occurrence of one jump within the previous 7 trading hours, and two jumps within the previous 14 trading hours. The indicators are properly scaled to achieve convergence in the maximum likelihood optimization procedure. These covariates are assumed to be related to non-homogeneous Poisson jump arrivals in an affine function: $d\Lambda_\theta(t) = \theta_0 + \theta_{1,j} dX_j(t)$ with $j = 1, 2$. The significance level α for Stage I is set equal to 5%. These results are based on 15- minute NYSE-transaction return data over the 5-year period, July 1, 2001 to June 30, 2006.

Table 8: **Jump and Analyst Recommendation in Individual Component Stocks of the DJIA Index**†

Ticker	$\hat{\theta}_0$	$\hat{\theta}_3$	JPT with $X_3(t)$	P-value
AA	0.0018	0.0386	14.2413	1.6080e-004
AIG	0.0018	0.0216	6.0867	0.0136
AXP	0.0020	0.0315	9.1264	0.0025
BA	0.0023	0.0242	5.9743	0.0145
C	0.0022	0.0262	6.9978	0.0082
CAT	0.0017	0.0353	8.4917	0.0036
DD	0.0019	0.0474	13.6086	2.2515e-004
DIS	0.0018	0.0506	18.6597	1.5625e-005
GE	0.0021	0.0411	10.4136	0.0013
GM	0.0023	0.0650	23.0084	1.6129e-006
HD	0.0020	0.0350	11.1716	8.3060e-004
HON	0.0020	0.0324	8.2682	0.0040
HPQ	0.0021	0.0368	13.0173	3.0863e-004
IBM	0.0019	0.0665	24.3490	8.0368e-007
JNJ	0.0024	0.0337	9.9618	0.0016
JPM	0.0022	0.0356	10.1142	0.0015
KO	0.0014	0.0423	11.7735	6.0080e-004
MCD	0.0022	0.0439	14.2317	1.6162e-004
MMM	0.0011	0.0465	9.0026	0.0027
MO	0.0026	0.0458	7.5146	0.0061
MRK	0.0024	0.0549	21.4373	3.6558e-006
PFE	0.0028	0.0265	8.4274	0.0037
PG	0.0015	0.0591	14.1606	1.6785e-004
T	0.0024	0.0305	15.0914	1.0243e-004
UTX	0.0018	0.0342	7.4819	0.0062
VZ	0.0017	0.0316	15.7951	7.0586e-005
WMT	0.0017	0.0353	13.2364	2.7457e-004
XOM	0.0017	0.0140	0.0140	0.0450

† This table contains the results for testing whether publication of analyst recommendations on equities is a jump predictor for the corresponding equity returns. $dX_3(t)$ is the indicator for time of publication of the research report (to 1 minute) in First Call — Analyst Recommendations History Database. These covariates are assumed to be related to non-homogeneous Poisson jump arrivals in an affine function: $d\Lambda_\theta(t) = \theta_0 + \theta_3 dX_3(t)$. The significance level α for Stage I is set equal to 5%. These results are based on 15-minute NYSE transaction return data over the 5-year period, July 1, 2001 to June 30, 2006.

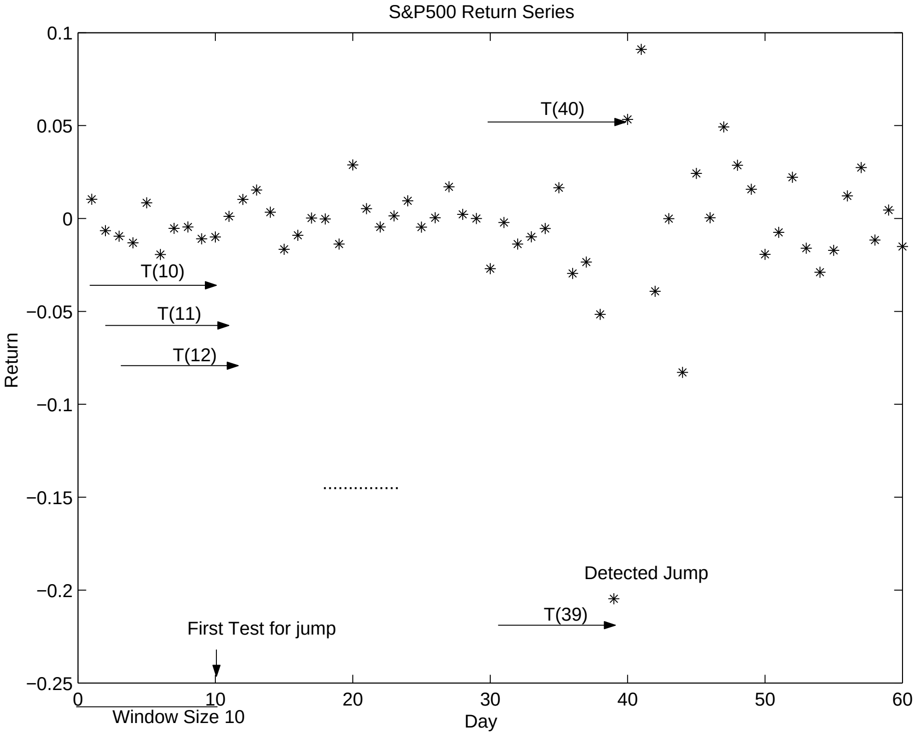
Table 9: **99%-VaR of Equity Return and Put Options**†

Daily	for Underlying		for In the Money		for Out of the Money	
Intensity	Static	Dynamic	Static	Dynamic	Static	Dynamic
	1.26	1.31	1.48	1.81	2.23	2.63
Difference(%)	3.97		22.29		17.94	
6 Hourly	for Underlying		for In the Money		for Out of the Money	
Intensity	Static	Dynamic	Static	Dynamic	Static	Dynamic
	0.24	0.43	0.38	0.69	0.56	0.94
Difference(%)	79.16		81.57		67.85	
3 Hourly	for Underlying		for In the Money		for Out of the Money	
Intensity	Static	Dynamic	Static	Dynamic	Static	Dynamic
	0.12	0.19	0.28	0.66	0.19	0.50
Difference(%)	58.33		135.71		163.15	

† This table reports the empirical 99%-VaR of equity returns and their put options, based on simulation. VaR values are all expressed as percentages. The underlying rate of return process has volatility at 30%, standard deviation of jump size distribution at 60%, and fixed jump intensity of 10 per year. The θ_1 coefficient for jump clustering is 5%.

Figure 1: **Intuition of Jump Predictor Test (JPT)**

Panel A: Stage I by jump detection in Lee and Mykland (2006)



Panel B: Stage II by the predictor test through partial likelihood