ECONSTOR
Make Your Publications Visible.

A Service of

ZBW Leibniz-Informationszentrum Wirtschaft
Leibniz Information Centre for Economics

Duong, Diep; Swanson, Norman

Working Paper

Volatility in discrete and continuous time models: A survey with new evidence on large and small jumps

Working Paper, No. 2011-17

Provided in Cooperation with:
Department of Economics, Rutgers University

Suggested Citation: Duong, Diep; Swanson, Norman (2011) : Volatility in discrete and continuous time models: A survey with new evidence on large and small jumps, Working Paper, No. 2011-17, Rutgers University, Department of Economics, New Brunswick, NJ

This Version is available at:
https://hdl.handle.net/10419/59488

Standard-Nutzungsbedingungen:

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.

Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

Terms of use:

Documents in EconStor may be saved and copied for your personal and scholarly purposes.

You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.

If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.

WWW.ECONSTOR.EU

Mitglied der Leibniz-Gemeinschaft

Volatility in Discrete and Continuous Time Models: A Survey with New Evidence on Large and Small Jumps*

Diep Duong and Norman R. Swanson*
Rutgers University

October 2010

The topic of volatility measurement and estimation is central to financial and more generally time series econometrics. In this paper, we begin by surveying models of volatility, both discrete and continuous, and then we summarize some selected empirical findings from the literature. In particular, in the first sections of this paper, we discuss important developments in volatility models, with focus on time varying and stochastic volatility as well as nonparametric volatility estimation. The models discussed share the common feature that volatilities are unobserved, and belong to the class of missing variables. We then provide empirical evidence on "small" and "large" jumps from the perspective of their contribution to overall realized variation, using high frequency price return data on 25 stocks in the DOW 30. Our "small" and "large" jump variations are constructed at three truncation levels, using extant methodology of Barndorff-Nielsen and Shephard (2006), Andersen, Bollerslev and Diebold (2007) and Aït-Sahalia and Jacod (2009a,b,c). Evidence of jumps is found in around 22.8% of the days during the 1993-2000 period, much higher than the corresponding figure of 9.4% during the 2001-2008 period. While the overall role of jumps is lessening, the role of large jumps has not decreased, and indeed, the relative role of large jumps, as a proportion of overall jumps has actually increased in the 2000s.

Keywords: Itô semi-martingale, realized volatility, jumps, quadratic volatility, multipower variation, tripower variation, truncated power variation, quarticity, infinite activity jumps.

* Diep Duong, Department of Economics, Rutgers University, 75 Hamilton Street, New Brunswick, NJ 08901, USA, dduong@econ.rutgers.edu. Norman R. Swanson, Department of Economics, Rutgers University, 75 Hamilton Street, New Brunswick, NJ 08901, USA, nswanson@econ.rutgers.edu. This paper has been prepared for the 9th Annual Advances in Econometrics Conference on Missing-data Methods held at the Southern Methodist University campus in Dallas in fall 2010. The authors owe a great deal of thanks to the organizers, and in particular to David Drukker, for many useful comments made during the writing of this paper. Diep and Swanson would like to thank the Research Council at Rutgers University for research support.

1 Introduction

In this paper, we begin by surveying models of volatility, both discrete and continuous, and then we summarize some selected empirical findings from the literature. In particular, in the first sections of this paper, we discuss important developments in volatility models, with focus on time varying and stochastic volatility as well as the nonparametric volatility estimation. The models discussed share the common feature that volatilities are unobserved, and belong to the class of missing variables. We begin by looking at the key ARCH class of models, followed by a discussion of the class of continuous time processes frequently used in finance and the link between discrete time and continuous time models. We then discuss the construction of implied volatility in the Black-Scholes framework, and generalizations thereof. and it's extensions. Finally, we discuss recent research in the area of "model free" estimation of integrated volatility via use of so-called realized volatility, and variants thereof called realized measures. In our empirical investigation, we use realized measures to investigate the role of jumps in the realized variation of stock price returns.

The importance of realized volatility to econometric modelling is now obvious. For example, future realized volatilities are often used in the so-called variance swap, an important product in the volatility derivative market. Other products that use realized volatility such as caps on the variance swap, corridor variance swaps, and options on realized volatility have also been introduced into the class of volatility related financial instruments traded in financial markets. The key here is that investors worry about future volatility risk, and hence often choose to opt for this type of contract in order to hedge against it.[1] Realized volatility is also needed for calculation of the variance risk premium, a new financial variable that has interesting implications in asset pricing. For example, Bollerslev, Tauchen and Zhou (2009) find that the variance risk premium is able to explain time-series variation in post-1990 aggregate stock market returns with high (low) premia predicting high (low) future returns. Finally, note that in the context of realized volatility, jumps have a significant impact on modeling and forecasting volatility and it's realized measures. For example, when jumps are present, realized volatility is a biased estimator of integrated volatility. Thus, practitioners who are interested in modeling risks associated with continuous components

[1] Volatility and variance swaps are newer hedging instruments, adding to the traditional volatility "Vega", which is derived from options data. See Hull and White (1997, pp. 328) for a definition of Vega. For example, as noted in Carr and Lee (2008), the UBS book was short many millions of vega in 1993, and they were the first to use variance swaps and options on realized volatility to hedge against volatility risk.

of return processes, or integrated volatility, should use carefully designed realized measures that take jumps effects into account.[2] Careful analysis of jumps and realized measures in the presence of jumps are crucial elements to any reasonable quantification of risk. Moreover, several authors[3] have found that separation of continuous components from jump components can improve forecasts of future realized volatility. This finding should be of great interest to practitioners, especially when their objective is hedging. In summary, risk, or volatility plays an important role in many areas of econometrics used in the finance industry. However, as volatility is generally unmeasured, it poses a standard sort of "missing variable" problem.

Turning back to our discussion of jumps, note that evidence of jumps in financial markets is plentiful. In an important paper, Huang and Tauchen (2005) find evidence of jumps for S&P cash and future (log) returns from 1997 to 2002, in approximately 7% of the trading days. Their test for jumps requires the jump component to be a compound Poisson process. Several authors, including Cont and Mancini (2007), Tauchen and Todorov (2008) and Aït-Sahalia and Jacod (2009c) have taken the analysis of jumps one step further by developing tests to ascertain whether the process describing an asset contains "infinite activity jumps" - those jumps that are tiny and look similar to continuous movements, but whose contribution to the jump risk of the process is not negligible. Cont and Mancini (2007) implement their method of testing for the existence of infinite activity jumps using foreign exchange rate data, and find no evidence infinite activity jumps. Aït-Sahalia and Jacod (2009c) estimate that the degree of activity of jumps in Intel and Microsoft log returns is approximately 1.6, which implies evidence of infinite activity jumps for these, and possibly many other stocks. Andersen, Bollerslev and Diebold (ABD: 2007) find that separating out the volatility jump component results in improved out-of-sample volatility forecasting, and find that jumps are closely related to macroeconomic announcements. In summary, it is now generally accepted that many return processes contain jumps.

In the part of this paper that present empirical findings relevant to the topic discussed in previous sections, we examine high frequency data for 25 stocks in the DOW 30, using 5 minute interval observations, and for the sample period from 1993 to 2008. Some of the stocks in our data set, (e.g. Microsoft and Intel) have been found to be characterized by infinite activity jumps by Aït-Sahalia and Jacod (2009b,c), and therefore do not belong to the class of finite activity jump

[2] See Corradi, Distaso and Swanson (2009, 2010) for discussion of prediction of integrated volatility.

[3] For instance, see Andersen, Bollerslev and Andersen (2007).

processes that Barndorff-Nielsen and Shephard (BNS: 2006) has often been applied to. This fact underscores the importance of the recent papers by Jacod (2008), Tauchen and Todorov (2008) and Aït-Sahalia and Jacod (2009a,b,c), where new limit theory applicable to infinite activity is implemented and developed; and underscores why the results of these papers are used in our empirical investigation. In summary, we find evidence of jumps in around 22.8% of the days in the 1993-2000 period, and 9.4% in the 2001-2008 period. This degree of jump activity implies more (jump induced) turbulence in financial markets in the previous decade than the current decade. However, and as expected, the prevalence of "large" jumps varies across these periods. (Note that we examine large jumps by picking 3 different fixed γ levels, corresponding to $50th, 75th$ and $90th$ percentiles of samples of the monthly maximum return increments, i.e. our monthly "abnormal event" samples.) In particular, large jump activity increases markedly during the 2001-2008 period, with respect to its contribution to the realized variation of jumps and with respect to the contribution of large jumps to the total variation of the (log) price process. This suggests that while the overall role of jumps is lessening, the role of large jumps has not decreased, and indeed, the relative role of large jumps, as a proportion of overall jumps has actually increased in the 2000s. Note that this result holds on average across all 25 stocks examined. In summary, it appears that frequent "small" jumps of the 1990s have been replaced with relatively infrequent "large" jumps in recent years. Interestingly, this result holds for all of the stocks that we examine, supporting the notion that their is strong co-movement across jump components for a wide variety of stocks, as discussed in Bollerslev, Law and Tauchen (2008).

The rest of the paper is organized as follows: Section 2 discusses volatility models in discrete time. Section 3 discusses volatility models in continuous time, and outlines various measures used to estimate (and forecast) volatility. Section 4 summarizes results from the extant testing and prediction literatures that are often used in the study of volatility, and in particular that are used in our empirical investigation that is presented in Section 5. Concluding remarks are contained Section 6.

2 Volatility Models in Discrete Time

Though the focus in this paper mainly involves continuous time models, we start by highlighting stylized facts and some key discrete time volatility models. The rationale for these highlights is not

only because discrete time models have a pivotal role in volatility research, but also because their set-up is intuitive and gives us interesting insights into the way econometricians build empirical facts into their models. We summarize some key results for the ARCH, GARCH and EARCH models of Engle (1982), Bollerslev (1986) and Nelson (1991). We finish by outlining the work of by Nelson (1990) and Corradi (2000)[4] that clarifies the link between discrete time and continuous time models. For a complete survey of conditional heteroskedasticity models, refer to Bollerslev, Chou and Kronner (1992) and Bollerslev, Engle and Nelson (1994).

2.1 Stylized Facts in Financial Market - Directions for Volatility Models

It is well-known in empirical finance that asset returns share various regularities, all of which guide financial economists and econometricians in their choice of models. These stylized facts have been discussed by many authors. Here, we highlight some of them that pertain to stock returns.

Leptorkurtosis: Asset returns have been noted by Mandelbrot (1963) and Farma (1965) to have fat tails and one therefore should use non-normal distributions to model their dynamics. Farma (1965) shows evidence of excess kurtosis in the distribution of stock returns. According to Clark (1973), a stochastic process is fat tailed if it is conditionally normal with a randomly changing conditional variance. Engle and González-Rivera (1991) introduce a semi-parametric volatility model, which allows for generic return distributions.

Volatility Clustering and Persistence: By observing cotton prices, Mandelbrot (1963) stressed that ".... large changes tend to be followed by large changes, of either sign, and small changes tend to be followed by small changes...". The persistence of shocks to the conditional variance of stock returns seems to be clear. The interpretation of this persistence as well as how long the shocks persist are crucial in specifying the "correct" dynamics. Porterba and Summers (1986) note that volatility shocks may affect the entire term structure, associated risk premia, and investment in long-lived capital goods.

Volatility persistence is an important feature that pertains to models with time varying and codependent variance. Black and Scholes (1973) wrote that "...there is evidence of non-stationary in the variance. More work must be done to predict variances using the information available.". Since their paper, numerous autoregressive conditional heteroskedasticity, volatility and stochastic

[4]There are several other papers in this line of research. For instance see Fornari and Mele (2006), among others.

volatility models have been developed.

Leverage Effects: Black (1976) observes that changes in stock prices seem to be negatively correlated with changes in stock volatility. Volatility seems to increase after bad news and decrease after good news. Schwert (1989, 1990) presents empirical evidence that stock volatility is higher during recessions and financial crises. Christie (1982) discusses economic mechanisms that explain this effect. Specifically, reductions in equity value raise the riskiness of firms, as implied by debt to equity ratios, and therefore lead to increases in future volatility. For modeling, Nelson (1991) suggests a new model that captures the asymmetric relation between returns and changes in volatility.

Co-movement in Volatilities. This is also first commented on by Black (1976). He pointsout the commonality in volatility changes across stocks. When stock volatilities change, they all tend to change in the same direction. This suggests that (few) common (unobserved or missing) factors might be specified when modelling individual asset return volatility.

2.2 ARCH and GARCH Models

Modeling and forecasting financial asset return volatility constitutes two of the cornerstones in modern econometric research. Much of the impetus for these lines of research owes to the seminal work of Engle (1982), in which he makes a key contribution to literature by introducing the autoregressive conditional heteroskedasticity (ARCH) model. Later, the Generalized ARCH (GARCH) model is developed by Bollerslev (1986), and these two models are broadly considered as the first to successfully capture the dependency of conditional variance on past information. In the ARCH model, conditional variances are simply modeled as a linear function of past squared error terms.

Uncovering the correct conditional volatility specification turns out to be important, especially in finance, where the central idea is the risk return trade-off. For investment decision-making, risk averse investors take into account not only expected return but also the level of risk. Investors demand risk premia for risk. This risk premium may include premia due to changes in volatility.

In risk neutral pricing, volatility is needed for the derivation of the market price of risk. ARCH models and its relatives offer convenient parametric specifications of the heterogeneous volatility process that is needed in this context. As pointed out in Engle (1982), the idea of time varying volatility had already been discussed in the early work of Granger and Andersen (1978), in the context of their so-called bilinear model. However, little evidence of these bilinear models has been

found in the data.

Turning now to our discussion of ARCH type models, let X_t be a financial asset return, say, and F_{t-1} denotes a filtration of all information through time $t-1$. The prototypical autoregressive conditional heteroskedasticity (ARCH) model has:

$$X_t = \varepsilon_t \sigma_t$$

$$\varepsilon_t \sim i.i.d \text{ with}$$

$$E(\varepsilon_t) = 0 \text{ and } Var(\varepsilon_t) = 1$$

$$\sigma_t^2 = \alpha_0 + \alpha_1 X_{t-1}^2$$

and in the more general case:

$$X_t | F_{t-1} \sim N(Z_t \beta, \sigma_t^2)$$

$$\sigma_t^2 = h(\varepsilon_{t-1}, \varepsilon_{t-2,\ldots,}\varepsilon_{t-p}, \alpha)$$

$$\varepsilon_t = X_t - Z_t \beta$$

where Z_t may contain lags of X_t. If the function h contains current and lagged X's, then

$$\sigma_t^2 = h(\varepsilon_{t-1}, \varepsilon_{t-2,\ldots,}\varepsilon_{t-p}, x_t, x_{t-1,\ldots,}x_{t-p}, \alpha)$$

In this class of models, ARCH(p) is the most popular where

$$X_t | F_{t-1} \sim N(Z_t \beta, \sigma_t^2)$$

$$\sigma_t^2 = \alpha_0 + \alpha_1 \varepsilon_{t-1}^2 + \alpha_2 \varepsilon_{t-2}^2 + \ldots \alpha_p \varepsilon_{t-p}^2$$

Engle (1982) proposes a convenient estimation and testing methodology for the model using maximum likelihood. He shows that α and β could be estimated separately under some regularity conditions[5]. To capture the trade-off between risk and expected return, Engle, Lilien and Robins (1987) introduce ARCH in mean, or ARCH-M models. Let

$$X_t = g(Z_{t-1}, \sigma_t^2; b) + \varepsilon_t$$

The appealing feature of this model is that the conditional mean, μ_t, is a function of the variance, i.e. $\mu_t = g(Z_{t-1}, \sigma_t^2; b)$. This helps us to model directly the risk-return relationship, and has important

[5]For details, see Sections 4 and 5 in Engle (1982).

implications for predicting the conditional mean function, since the conditional volatility enters therein. The parametric choice for the function g depends on the modeler. In practice, many papers set g to be a linear or logarithmic function.

An important improvement to these models is made by Bollerslev (1986), where the ARCH model is generalized to the Generalized ARCH (GARCH) model. As noted in Bollerslev (1986), the extension from ARCH to GARCH is similar to the extension in time series modelling of an AR to an ARMA model. Specifically, as in the case of the ARCH model, let ε_t be the innovation in a linear regression

$$\varepsilon_t = X_t - Z_t'\beta$$

Then the GARCH (p,q) specification is given by

$$\varepsilon_t | F_{t-1} \sim N(0, \sigma_t^2)$$

$$\sigma_t^2 = \alpha_0 + \sum_{i=1}^{q} \alpha_i \varepsilon_{t-i}^2 + \sum_{i=1}^{p} \beta_i \sigma_{t-i}^2$$

$$\varepsilon_t = X_t - Z_t\beta$$

where p and q denote lag orders, and

$$p \geq 0, q > 0$$

$$\alpha_0 > 0,\ \alpha_i \geq 0,\ i = 1, ..., q$$

$$\beta_i \geq 0,\ i = 1, ..., p$$

It is clear that the difference between the above set up and the ARCH model is the linear lagged conditional variances. Conditional volatility today not only depends on the lagged innovations but also on lagged conditional volatilities.

Bollerslev (1986) presents a complete set of results on the conditions under which the model is appropriate, as well as maximum likelihood and testing procedures for implementing the general GARCH (p,q) model. The most successful model, empirically, is the GARCH(1,1) model. Engle and Bollerslev (1986) discussthe so-called Integrated GARCH or IGARCH model. Under this specification, $\sum_{i=1}^{q} \alpha_i + \sum_{i=1}^{p} \beta_i = 1$, and this leads to a unit root in the volatility equation.

In other key papers, Nelson (1990, 1991) discusses the use of EARCH (i.e., exponential ARCH) to approximate continuous time processes. Nelson (1991) points out that the GARCH model has

several limitations in empirical applications to financial markets. For instance, in the GARCH model, volatility responds symmetrically to positive and negative residuals and therefore does not explain the stylized leverage effect. In lieu of this, Nelson (1991) proposes the Exponential ARCH, or EARCH in which the volatility function is constructed as follows:

$$X_t = \sigma_t \varepsilon_t$$

$$\varepsilon_t \ \sim i.i.d \text{ with}$$

$$E(\varepsilon_t) = 0 \text{ and } Var(\varepsilon_t) = 1$$

and

$$\ln(\sigma_t^2) = \alpha_t + \sum_{k=1}^{\infty} \beta_k g(\varepsilon_{t-k}), \qquad \beta_1 \equiv 1$$

where $\{\alpha_t\}_{t=-\infty,\infty}$ and $\{\beta_k\}_{k=1,\infty}$ are parameters.

The choice for the functional form of g is

$$g(\varepsilon_t) = \theta\varepsilon_t + \gamma(|\varepsilon_t| - E|\varepsilon_t|)$$

This set-up allows the conditional variance process to respond asymmetrically to rises and falls in stock prices. It is straightforward to verify this as when ε_t is positive $g(\varepsilon_t) = (\theta+\gamma)\varepsilon_t - \gamma E(|\varepsilon_t|)$ and when ε_t is negative $g(\varepsilon_t) = (\theta-\gamma)\varepsilon_t - \gamma E(|\varepsilon_t|)$. In each case, the $g(\varepsilon_t)$ is linear function with a different slope.

In addition, Nelson (1991) points out, while in GARCH, it is difficult to verify the persistence the shocks to the variance, in the EARCH model, the stationarity and ergodicity of the logarithm of the variance process are easily checked. He states conditions for the ergodicity and strict stationarity of $\{exp(-\alpha_t\sigma_t^2)\}$ and $\{exp(-\alpha_t/2X_t)\}$.[6] Other modifications of the GARCH (1,1) model include the GJR model proposed by Glosten, Jaganathan and Runkle (1993). This model imposes structure that induces asymmetry in shocks to returns in a different way. Namely, they define

$$\sigma_t^2 = \omega + \alpha\varepsilon_t^2 + \gamma\varepsilon_t^2 1_{\{\varepsilon_t \geq 0\}} + \beta\sigma_{t-1}^2$$

Note that when $\gamma < 0$, positive return shocks increase volatility less than negative shocks.

The above discussion summarizes a very few of the important models in class of discrete ARCH models.[7] In addition to these models, there have been many modifications and improvements. For

[6] See Theorem 2.1 in Nelson (1991).

[7] We also present the work of Heson and Nandi (2000) and Barone-Adesi, Engle, Mancini (2008) which use modifications of ARCH models for option pricing in Section 3.2.2.

a complete list and discussion, see Bollerslev (2008), where he provides a Glossary to ARCH. For models with multivariate specifications (see Bollerslev, Engle and Wooldbridge (1988)). In the next section, we will highlight some links between discrete time models and continuous time models in the framework of modeling volatility.

2.3 From ARCH and GARCH to Continuous Time Models

An interesting aspect of the volatility literature is the connection between discrete time and continuous time models. In the case of constant volatility, the classical result by Cox and Ross (1976) shows that the limiting form of the jump process

$$dX_t = \mu X_t dt + cX_t dN_t(\lambda)$$

as $\lambda \to 0$ is the diffusion process

$$dX_t = \mu X_t dt + \sigma X_t dW_t$$

where σ is a function of c. $N_t(\lambda)$ is a continuous time Poisson process with intensity λ, (i.e., dN_t is the number of jumps of X_t during dt and is Poisson-distributed with parameter λdt). cX_t is the jump amplitude and W_t is a standard Brownian motion. In another important research, Nelson (1990) bridges the gap between discrete and continuous time stochastic volatility models by using AR(1) Exponential ARCH and GARCH (1,1) models as approximations for continuous time processes. It should be noted that in this approximation framework, only the discrete models with one lag are relevant due to the characteristics of continuous time models. GARCH models with two more lags as explanatory variables are not relevant. Under certain assumptions [8], Nelson (1990) at first looks at the GARCH (1,1) - M process of Engle and Bollerslev (1986)

$$\begin{aligned} X_t &= X_{t-1} + c\sigma_t^2 + \sigma_t \varepsilon_t, \\ \sigma_t^2 &= \omega + \sigma_t^2[\beta + \alpha \varepsilon_t^2], \end{aligned}$$

If time is partitioned more finely, one can write the above difference equation as

$$\begin{aligned} X_{kh} &= X_{(k-1)h} + hc\sigma_{kh}^2 + \sigma_{kh}\varepsilon_{kh} \\ \sigma_{(k+1)h}^2 &= \omega_h + \sigma_{kh}^2[\beta_h + h^{-1}\alpha_h \varepsilon_{kh}^2], \end{aligned}$$

[8] For details, see Section 2 in Nelson (1990).

where h is the time increment and $\varepsilon_{kh} \sim i.i.d\ N(0,h)$. He shows that if h goes to 0 in the limit, this system converges weakly in distribution[9] to a diffusion process of the form

$$dX_t = c\sigma_t^2 dt + \sigma_t dW_{1t}$$
$$d\sigma_t^2 = (\omega - \theta\sigma_t^2)dt + \alpha\sigma_t^2 dW_{2t}$$

where W_{1t} and W_{2t} are linearly independent standard Brownian motions, independent of the initial values (X_0, σ_0^2).

In another important paper in this line of research, Corradi (2000) considers the limit when $h \to 0$, of the GARCH (1,1) process

$$X_{kh} - X_{(k-1)h} = \sigma_{(k-1)h}\varepsilon_{kh},$$

$$\sigma_{kh}^2 - \sigma_{(k-1)h}^2 = \omega_h + (\omega_{1h} - 1)\sigma_{(k-1)h}^2 + h^{-1}\omega_{2h}\sigma_{(k-1)h}^2\varepsilon_{kh}$$

She shows that in the limit, this system converges to either one of the following continuous time processes, depending on the parameters ω_{1h} and w_{2h} [10]

$$dY_t = \sigma_t dW_t$$
$$d\sigma_t^2 = (\omega_0 + \theta\sigma_t^2)dt$$

or

$$dY_t = \sigma_t dW_{1t}$$
$$d\sigma_t^2 = (\omega_0 + \theta\sigma_t^2)dt + \alpha\sigma_t^2 dW_{2t}$$

where (W_{1t}, W_{2t}) are two standard independent Brownian motions.

Nelson (1990) also introduces a class of ARCH models which can approximate a wide range of stochastic differential equations. He investigates the approximations of the system of stochastic differential equations defined as follows:

$$dX_t = f dt + g dW_{1t}$$
$$dY_t = F dt + G dW_{2t}$$
$$\begin{bmatrix} dW_{1,t} \\ dW_{2,t} \end{bmatrix} [dW_{1,t\ 2,t}] = \Omega dt$$

[9] For a definition of weak convergence for stochastic processes, see Billingsley (1978).

[10] For details, see Proposition 2.1 in the Corradi (2000).

where the first equation is univariate, Y_t is vector of latent state variables, W_{1t} can be correlated with elements in $W_{2t.}$, and F, G, f, g are functions of X_t, Y_t, and t. He shows that the above system is the limit of the following ARCH type discrete time system of difference equations

$$X_{(k+1)h} = X_{kh} + fh + g\varepsilon_{kh}$$
$$Y_{(k+1)h} = Y_{kh} + Fh + G\varepsilon^*_{kh}$$

where ε^*_{kh} corresponds to W_{2t} and is constructed on the basis of $\varepsilon_{kh} \sim N(0, h)$ and Ω.[11]

3 Volatility in Continuous Time

3.1 Continuous Time Models

Continuous time models are a natural choice for describing the evolution of term structures, stock prices and other financial variables, and are integral for the pricing of financial instruments that depend on such variables. In this class of models, the choice of volatility function is a key issue in model selection. This choice is generally both theoretically and empirically driven. As an example of their importance, note that over the last 20 years, continuous time models have taken on a central role in option pricing, risk management and volatility forecasting. One key advantage of this class of models is that it lends itself naturally to the use of recently made available high frequency data. In principle, financial managers can update their dynamic trading strategies every second. For econometrician, the use of high frequency data has interesting implications for both estimation and prediction.

In previous section, under the framework of time varying and stochastic volatility, we already mentioned some results that use discrete time ARCH and GARCH models as inputs to approximate diffusion processes in continuous time. In this section we summarize some standard continuous time models that have been used in asset pricing as well as term structure modelling.

Though this is not meant to be an exhaustive list, we outline various standard continuous time models, from simple to more complicated. Note that we first focus on the specification of the single equation (solution) models[12]. Also note that, in each specification, volatility is modeled differently:

[11] For details on the construction of ε^*_{kh}, see Section 3.2 in Nelson (1990)

[12] For a definition of diffusion processes and stochastic differential equations, see Karlin and Taylor (1981) pp. 157 and 373, respectively.

Diffusion Processes:

Brownian Motion with Drift:

$$dX_t = \mu dt + \sigma dW_t$$

This specification has been used a lot in early work in economics and finance due to its simplicity. It is obvious that X_t is normally distributed with mean μt and variance $\sigma^2 t$.

Geometric Brownian Motion (Log Normal Model):

$$dX_t = \mu X_t dt + \sigma_t X_t dW_t$$

This model has been very popular for asset prices. It has been extensively used in the Black and Scholes (1973) option pricing framework and in structured corporate finance. The main drawback of this model is that the (log) return process has constant volatility. To see this, apply Itô' Lemma[13] to the function $f(x) = \log(x)$, yielding

$$d\log(X_t) = (\mu - \frac{\sigma^2}{2})dt + \sigma dW_t.$$

Ornstein-Uhlenbeck Process (sometimes referred to as the Vasicek (1977) model, and often used to model interest rates):

$$dX_t = (\alpha + \beta X_t)dt + \sigma dW_t.$$

Cox, Ingersoll and Ross (1985) analyze the following square root process, also known as the CIR model, again for modelling the term structure of interest rates, although this model, unlike the Vasicek model, ensures positivity of rates.

$$dX_t = (\alpha + \beta X_t)dt + \sigma\sqrt{X_t}dW_t.$$

Brennan and Schwartz (1979) and Courtadon (1982) analyze the process:

$$dX_t = (\alpha + \beta X_t)dt + \sigma X_t^2 dW_t.$$

[13] Itô' Lemma formula is presented at the end of this section.

Constant Elasticity of Variance (CEV):

$$dX_t = \mu X_t dt + \sigma X_t^{\beta/2} dW_t$$

Note that the interpretation of this model depends on β. In particular, in the case using this process to model stock price, if $\beta = 2$, the price process X_t follows geometric Brownian motion and therefore the volatility of the (log) stock return process is constant. If $\beta < 2$, this model captures the leverage effects discussed above. Among other authors, Beckers (1980) uses this CEV model for stocks. Marsh and Rosenfeld (1983) apply the CEV model to interest rates and Emanuel and Macbeth (1982) use this set-up for option pricing.

Generalized Constant Elasticity of Variance:

$$dX_t = (\alpha X_t^{-(1-\omega)} + \beta X_t)dt + \sigma X_t^{\beta/2} dW_t$$

This process nests the log diffusion when $\beta = 2$, and square root process when $\beta = 1$.

Chan, Karolyi, Longstaff, and Sanders (1992) study the case of linear drift and CEV diffusion with $\beta \geq 2$

$$dX_t = (\alpha + \beta X_t)dt + \sigma X_t^{\beta/2} dW_t.$$

Duffie and Kan (1996) specify a mean reversion and square root structure in volatility for modelling of interest rates. In the univariate case:

$$dX_t = (\alpha - X_t)dt + \sqrt{\sigma_0 + \sigma_1 X_t} dW_t.$$

Aït-Sahalia (1996) looks at more the general case of diffusions with general drift and CEV:

$$dX_t = (\alpha + \beta X_t + \gamma X_t^2 + \eta / X_t)dt + \sigma X_t^{\beta/2} dW_t.$$

Needless to say, model selection is important issue when specifying diffusion models. Note that in the general setting, the diffusion process is written as

$$dX_t = \mu(X_t, t, \theta)dt + \sigma(X_t, t, \theta)dW_t$$

Also, it is known that the drift and diffusion terms $\mu(X_t, t)$ and $\sigma(X_t, t)$ respectively uniquely determine the stationary density, say $f(x, \theta_0)$, associated with the invariant probability measure of the above diffusion process [14]. In particular,

$$f(x, \theta_0) = \frac{c(\theta_0)}{\sigma^2(x, \theta_0)} \exp\left(\frac{2\mu(u, \theta_0)}{\sigma^2(u, \theta_0)} du\right)$$

In a seminal paper, Aït-Sahalia (1996) constructs a nonparametric test for interest rate models on the basis of the comparison of such stationary densities. In his empirical application to spot interest rates, he finds that the misspecification of the models in the literature on spot interest rates is mainly due to the linearity of the drift function in such models. In addition, his proposed model (general drift and CEV) could not be rejected. In this same line of research, Corradi and Swanson (2005, 2010) develop bootstrap specification tests for univariate and multifactor diffusion process that do not require knowledge of the transition density. Instead of comparing of densities, their method is based on a comparison of cumulative distribution functions. They also extend their methods to diffusion process with jumps and stochastic volatility.

Jump Diffusions:

Define a jump diffusion process as the combination of two components. The first component is the continuous process which is specified by a diffusion process and the second component is specified by a jumps process, specifically

$$dX_t = \mu(X_t, t, \theta)dt + \sigma(X_t, t, \theta)dW + dJ_t$$

where J_t is pure jump process which jointly depend on two source of randomness; namely the magnitude of the jump, say with distribution ν, and jump intensity, say $\lambda(X_t)$. One pioneering work which incorporates jumps into continuous time processes is Merton (1976), where he models the continuous component of the log price process to be Gaussian as in the case of geometric Brownian motion. The magnitude of jumps also follows a Gaussian distribution, and jumps follow Poison distribution in his paper. Newer developments in this area do not "append" a "discrete" jump process onto the diffusion, but instead specify the jumps using other means, such as via the use of Levy processes.

Affine Jump Diffusion Model: This class of models is widely studied in the empirical finance literature, especially in term structure modelling. The family of affine processes X_t including jumps

[14] See Karlin and Taylor (1981), pp. 241.

is parametrized as follow

$$dX_t = \kappa(\alpha - X_t)dt + \Omega\sqrt{D_t}dW + dJ_t$$

where W_t is an $N-$dimensional independent standard Brownian motion, κ and Ω are square $N \times N$ matrices. D_t is a diagonal matrix with ith diagonal element given by

$$d_{ii,t} = \theta_i + \delta_i' X_t$$

The jump intensity is assumed to be a positive, affine function of X_t and the jump size distribution is assumed to be determined by its conditional characteristic function. As shown by transform analysis in Duffie, Pan and Singleton (2000), the attractive feature of this class of affine jump diffusions is that the exponential affine structure characteristics function is known in closed form. Namely

$$\Phi_t(X_t) = \exp(a_t + b_t' X_t)$$

where functions a_t and b_t can be derived from Riccati equations.[15] With known characteristics function, one can use either GMM to estimate the parameters of this system of this jump diffusion, and can use quasi-maximum likelihood (QML), once the first two moments are obtained. In the univariate case without jumps, as a special case, this corresponds to the above general CIR model with jumps.

Stochastic Volatility Models:

Stochastic volatility models are popular, particularly for modelling asset prices and interest rates. They are first introduced by Harvey, Ruiz and Shephard (1994) in discrete time. Stochastic volatility implies that unobserved volatility follows another stochastic process. For example, one specification could be

$$dX_t = (\alpha + \beta X_t)dt + \sigma_t dW_{1t},$$

and the volatility process follows:

$$d\sigma_t^2 = \kappa(\vartheta - \sigma_t^2)dt + \delta\sigma_t dW_{2t},$$

where $Cov(dW_{1t}, dW_{2t}) = \rho dt$.

Andersen and Lund (1997) estimate the generalized CIR model with stochastic volatility:

$$dX_t = \kappa_1(\alpha - X_t)dt + \sigma_t X_t^{\beta} dW_{1t},$$

[15] For details, see Singleton (2006), pp. 102.

$$d\log\sigma_t^2 = \kappa_1(\alpha - \log\sigma_t^2)dt + \delta dW_{2t}.$$

Mixed Stochastic Volatility, Jump Diffusion Models. An example of these models comes from the application of spectral GMM in Chacko and Viceira (2003) where the return process is specified as:

$$\begin{aligned} dX_t &= (\mu - \frac{\sigma_t^2}{2})dt + \sigma_t dW_{1t} + [\exp(J_u) - 1]dN_u(\lambda_u) + [\exp(-J_d) - 1]dN_d(\lambda_d) \\ d\sigma_t^2 &= \kappa(\alpha - \sigma_t^2)dt + \delta\sigma_t dW_{2t}, \end{aligned}$$

where λ_u, λ_d are jump intensity parameters and are constant, and where. J_u and J_d >0 are stochastic jump magnitudes that follow an exponential distribution, i.e.

$$f(J_u) = \frac{1}{\eta_u}\exp(\frac{-J_u}{\eta_u}),$$

$$f(J_u) = \frac{1}{\eta_d}\exp(\frac{-J_d}{\eta_d}).$$

In addition, in the option pricing literature, many models are nested in the following data generating process which allows for jumps in both equations

$$\begin{aligned} dX_t &= \mu_t dt + \sigma_t X_t dW_{1t} + dJ_{1t} \\ d\sigma_t^2 &= \kappa(\alpha - \sigma_t^2)dt + \delta\sigma_t(\rho dW_{1t} + \sqrt{1-\rho^2}dW_{2t}) + dJ_{2t} \end{aligned}$$

where W_{1t} and W_{2t} are two independent Brownian motions process, and J_{1t} and J_{2t} are two jump processes. Popular models that are nested in this class include Heston (1993) with no jumps in either price or volatility, Bates (2000), Chernov and Ghysels (2000) and Pan (2002)[16].

Itô's Lemma Formula and Stochastic Calculus:

Earlier, we made use of this lemma. For completeness, then, we now provide a short summary of this key result in stochastic calculus. Itô' Lemma formula is crucial because it is essential for pricing financial derivatives. We will state a version of this lemma with respect to semi-martingale processes and will apply it to diffusion processes. For details see Protter (1990, p.71).

[16] For a more detailed discussion, see Singleton (2006), chapter 15.

Itô's Lemma: Let X be a semimartingale and let f be a continuous, twice differentiable function. Then $f(X)$ is a semimartingale, and the following formula holds

$$f(X_t) - f(X_0) = \int_0^t f'(X_{s-})dX_s + \frac{1}{2}\int_0^t f''(X_{s-})d[X,X]_s^c + \sum_{0\leq s\leq t}\left\{f(X_s) - f(X_{s-}) - f'(X_{s-})\Delta X_s\right\}$$

where $[X,X]_s^c$ is the quadratic variation process of the continuous component of X_t and ΔX_s is the size of the jump at the jump time s. Note that the first part of the formula is the standard result for continuous processes. The second part is the effect of jumps. Without jumps, the lemma applies to the process

$$dX_t = \mu(X_t,t)dt + \sigma(X_t,t)dW_t$$

Therefore $d\ [X,X]_s^c = \sigma^2(X_t,t)dt$ and the formula can be written as

$$f(X_t) - f(X_0) = \int_0^t (f'\mu + \frac{1}{2}f''\sigma^2)dt + \int_0^t f'\sigma dW_t$$

This is a standard and well-known formula used in describing continuous processes.

Finally, note that another fundamental tool in the finance literature is the change of measure, or Girsanov's theorem. Under risk neutral pricing, one needs to transform the models from the statistical measure P to risk neutral measure Q. Girsanov theorem provides the link between the two. For further details on the topic, refer to Duffie (2003, chapter 6).

3.2 Implied Volatility from Option Pricing

3.2.1 Black-Scholes Framework as an Illustration

Implied volatility is considered to be the market prediction of future volatility, used in the context of option pricing. In this section, we firstly provide a standard method to show how econometricians can construct implied volatility from Black-Scholes option prices. Though stochastic volatility better captures the dynamics of asset returns, BS is still considered to be an important element of option pricing theory and practice. However, that said, after the current discussion we turn to a discussion of "model free" measures of implied volatility, and the so-called VIX volatility index.

Within the Black-Scholes framework, we restate the derivation of the European call price. In this model, stock prices are assumed to be log-normally distributed. A nice feature of this assumption is that option prices can then be derived in closed form, i.e. using the Black-Scholes (BS) formula.

Once an option pricing function is known, implied volatility can in turn be backed out using stock, option and interest rate data. Note that one can also derive the price of any derivative whose payoff is a function of stock prices. Also, even under more complicated assumptions about the stock price process, the risk neutral pricing methodology presented below can still be applied.

Let's assume that we are interested in an asset market which has 3 assets: a riskless bond, a stock, and a derivative whose payoff is a function of the stock price. Also, say that under the physical probability measure, P, the price of a non-dividend paying stock, X_t, follows a geometric Brownian motion, i.e.

$$dX_s = \mu X_s ds + \sigma X_s dW_s$$

and the interest rate, r,associated with riskless bond (i.e. the short rate) is assumed to be constant. For simplicity, we analyze the dynamics of the process $\overline{X_s} = \frac{X_s}{ers}$. Using Itô' Lemma, $\overline{X_s}$ follows:

$$d\overline{X_s} = (\mu - r)\overline{X_s} ds + \sigma \overline{X_s} dW_s$$

Say that the call option on the stock has strike price K and maturity date T. It's payoff at time T is

$$C_T = (X_T - K)^+ = \max(X_T - K, 0)$$

Option pricing means we look for the price of the derivative c_t whose payoff at maturity is C_T. With the no arbitrage assumption, there must exist a risk neutral measure Q. Under this risk neutral measure, the discount process $\overline{X_s}$, $\overline{C_s} = \frac{C_s}{esr}$ has no drift term, or is a martingale:

$$\begin{aligned} d\overline{X_s} &= \sigma \overline{X_s} d\overline{W_s} \\ \overline{c_t} &= E^Q[\frac{C_T}{e^{rt}}|X_t] = e^{-rT} E^Q[((X_T - K)^+)|X_t] \end{aligned}$$

where $\overline{W_s}$ is a Brownian motion under measure Q. The expectation above is taken under the probability measure Q. Recall that probability measures P and Q can be transformed back and forth though Girsanov's Theorem. In particular $\overline{X_T}/\overline{X_t}$ is log-normally distributed with mean $-(\sigma^2/2)(T-t)$ and variance $\sigma^2(T-t)$. Or, given that the asset $\overline{C_s}$ has no drift term, one can easily verify that the c_t process is the solution to the following partial differential equation

$$-rf + f_t + rSf_X + 1/2\sigma^2 f_{XX} = 0$$

with the boundary condition at the expiration date $f(S,T)$ = Payoff at time $T = \max(X_T - K, 0)$. Note that one can use this derivation approach for any asset, with a change in payoff function. In

more complicated settings with more complicated price process dynamics, the same PDE approach can be applied.

However, in such cases, to solve the PDE, a numerical algorithm is needed as there may not be a closed form solution for the option price. Turning again to our closed form BS solution, note that it takes the form

$$c_t = \Phi(z) - e^{-r(T-t)}\Phi(z - \sigma\sqrt{T-t})$$

where Φ is the cumulative normal cdf and,

$$z = \frac{\log(X_t/K) + (r + \sigma^2/2(T-t))}{\sigma\sqrt{T-t}}$$

From the BS formula, one can invert the unobservable σ. In particular, σ is function of current time t variables, $X_t, r, \tau = T - t, K$ and c_t which are all observed and available in the data. This is useful in the framework of no-arbitrage pricing, and this option pricing tool is the key to backing out implied volatility. Generally, one can write

$$c_t = F(X_t, r, K, \tau, \sigma_\tau)$$

and the volatility process can be inferred once the nonlinear function F is known. If it is not known in closed form, we still can back out implied volatility via numerical analysis.

3.2.2 Deviation from Black-Scholes

The main drawback of Black-Scholes is that it is not consistent with empirical evidence that implied volatility varies across different maturities and strike price. Dumas, Flemning and Wahley (1998) propose an *ad hoc* Black-Scholes model in which volatility is not constant. This measure is now widely used by practitioners. *Ad hoc* BS allows for different implied volatilities to price options differently. In particular, implied volatilities are modeled as follows:

$$\sigma(\tau, K) = \alpha_0 + \alpha_1 K + \alpha_2 K^2 + \beta_0 \tau + \beta_1 \tau^2 + \beta_2 \tau K$$

where $\sigma(\tau, K)$ is the Black-Scholes implied volatility for and option with strike K and time to maturity τ. This method has been shown by Dumas et al. (1998) to be better than the constant volatility approach. Option pricing with varying volatility, such as a square root diffusion model, are also proposed by many researchers (for example, see the work of Bates (1996, 2000), Bakshi,

Cao and Chen (1997) and Scott (1997)). Heston (1993) provides a parametric stochastic volatility model, i.e. square root process for volatility, and solves for closed-form prices. Chernov and Ghysels (2000) and Pan (2002) introduces a more general pricing framework, i.e. stochastic volatility with jumps. For a survey of this option pricing literature, see Bates (2003), and for a discussion of the econometrics of option pricing see Garcia, Ghysels and Renault (2010). In an interesting discrete-time set-up, Heston and Nandi (2000) proposes a closed form GARCH option price model in which they present an option formula for a stochastic volatility model with GARCH. This discrete-time set-up is close to the Heston (1993) continuous time stochastic volatility model, but is easier to implement. They specify the following dynamics for asset returns

$$\log(X(_t) = \log(X_{t-h}) + r + \lambda\sigma_t^2 + \sigma_t\varepsilon_t$$
$$\sigma_t^2 = \omega + \beta_1\sigma_{t-h}^2 + \alpha_1(\varepsilon_{t-h} - \gamma_1\sigma_{t-h})^2$$

where α_1 determines the kurtosis of the distribution and γ_1 allows for asymmetric effects of shocks. Specifically

$$Cov_{t-h}(\sigma^2(t+h), \log(X_t)) = -2\alpha_1\gamma_1\sigma^2(t).$$

In testing the empirical implications of this GARCH option pricing model, they find that the model produces smaller valuation errors compared to the *ad hoc* BS model mentioned above. In different work, Barone-Adesi, Engle and Mancini (2008) propose a new option pricing method with filtered historical innovations. The new feature in their methodology is that it fits in an incomplete markets framework and is not based on the specification of the change of measure, i.e. from physical measure to risk neutral measure and state price density [17]. Instead, they estimate separate GARCH parameters in the risk neutral world. They show that their pricing outperforms other discrete GARCH models.

In recent papers in the semi-parametric literature, as opposed to Black-Scholes, Carr and Madan (1998), Demeterfi et al. (1999), Britten-Jones and Neuberger (2000), Lynch and Panigirtzoglou (2003), Jiang and Tian (2005), Car and Wu (2009) develop variants of so called "model-free" implied volatility. These estimators are referred to as semiparametric measures, as volatility is implied from option prices via risk neutral pricing without many of the usual parametric assumptions on the dynamics of asset returns. In addition, these volatility measures provide ex ante risk neutral

[17] For further discussion on risk neutral pricing and change of measures, see Duffie (2003), chapter 6.

expectations of future volatilities, which is an important input to calculate variance risk premia. Variance risk premia are defined as the difference between implied volatility and realized volatility. When (log) stock prices follow a continuous process, the implied volatility between time t and $t+h$ can be derived by the formula.

$$IMV_{t,t+h} = 2\int_0^\infty \frac{C(t+h,K) - C(t,K)}{K^2} dK = E_t^Q(\sigma_{t+h}^2)$$

where $C(t,K)$ is the price of European call option written on strike price K and maturing at time t. Here, $E_t^Q(\sigma_{t+h}^2)$ is the expectation of the variation of the log price process, or of the realized volatility. An advantage to using the above model free implied volatility in equity markets is that one now can rely on a published volatility index (usually the VIX) as a standard measure of implied volatility on the S&P 500 index. VIX is considered a key measure of market expectations of near-term volatility implied from S&P 500 stock index option prices. VIX was first introduced by the Chicago Board of Exchange (CBOE) in 1993, and often referred to by many as a "fear" index. In 2003, CBOE updated the calculation of VIX and the general formula is as follows:

$$IMV = \frac{2}{T}\sum_i \frac{\Delta K_i}{K_i^2} e^{r\tau} Q(K_i) - \frac{1}{\tau}\left[\frac{F}{K_0} - 1\right]^2$$

where $VIX = IMV * 100, \tau$ is time to expiration, F is the forward index level derived from index option prices, K_0 is the first strike price below F, K_i is the strike price of ith out of the money option, r is risk free rate, $\Delta K_i = (K_{i+1} - K_{i-1})/2$, $Q(K_i)$ is the mid point bid-ask spread for each option with strike price K_i.[18]

3.3 Realized Volatility - Nonparametric Measures

The latest developments in the volatility literature largely center on the use of so called realized volatility (RV) as a "model-free" estimator of latent variance of stock returns or other financial variables. Daily realized volatility is simply the sum squared returns from high-frequency data over short time interval within any given day. As noted in the key early works of Andersen and Bollerslev (1998), Banrdorff-Nielsen and Shephard (2002) and Medahhi (2002), RV and it's variants yield much more accurate ex post observations of volatility than the traditional sample variances

[18] For details, see "CBOE Volatility Index - VIX" at the link http://www.cboe.com/micro/vix/vixwhite.pdf and Demeterfi et al. (1999)

which used daily or lower frequency data. Many papers have been written on this topic since these first papers. In practice, RV has been an important variable in the volatility derivatives market. For instance, trading of forward contracts on future realized volatility was sporadically seen in financial market as early as 1993. This type of product is now common, and is often referred to as the variance swap. One important feature of this product is that it's payoff is a linear function of RV and therefore is simple to use as a hedging tool, compared to traditional vega hedging. Moreover, there is much market demand for this product as practitioners prefer implied volatility to variance and they need additional instruments to hedge against future volatility risk. Other products that used realized volatility such as caps on variance swaps, corridor variance swaps, and options on realized volatility are also available in financial markets. [19] In research, several authors have developed the concept of variance risk premia which directly depend on RV and they argue that this variable is useful in asset pricing. Variance risk premium (VRP) is defined as

$$VRP_t = IMV_t - RV_t$$

where IMV_t is implied volatility defined under the risk neutral measure Q as

$$IMV_t = E^Q[\text{Return variation between } t-1 \text{ and } t]$$

Bollerslev, Tauchen and Zhou (2009) use this premium to predict stock market returns and they find that the premium is able to explain time-series variation in post-1990 aggregate stock market returns with high (low) premia predicting high (low) future returns.

RV non-parametrically measures the variation of return processes, and the dynamics of RV can be driven by components other than those directly involved with returns. When the return process is continuous, it's variation is due to the continuous component, and is known as the integrated volatility (IV). Realized volatility is a proxy for IV[20]. Several authors (for example, Huang and Tauchen (2005) and Aït-Sahalia and Jacod (2009a,b,c)) find important evidence of active jumps in equity markets. If jumps occur, variation of the return process is greeater than integrated volatility as it contains a jump variation component. Realized volatility therefore is not an estimator of integrated volatility. ABD (2007) construct a simple measure of the variation of this

[19] For a discussion of volatility and variance swaps, see for instance, Carr and Lee (2009).

[20] For instance, under the assumption that the return process is continuous, Kristensen (2010) develops a kernel based approach to estimate integrated and spot volatilty using realized volatility.

jump component and then show that incorporation of jumps can affect estimation and volatility prediction. (Our empirical application discussed below fits within this strand of the literature. In particular, we provide evidence of jumps and large jumps as well as providing a measure of large jump variation in equity markets.)

Following the general set-up of Aït-Sahalia and Jacod (2009b), consider the filtered probability space $(\Omega, F, (F_t)_{t\geq 0}, P)$, in which $(F_t)_{t\geq 0}$ is denoted as a filtration (i.e., a family of sub-sigma algebras F_t of F, being increasing $t : F_s \subset F_t$ if $s \leq t$). The log price process, $X_t = log(P_t)$, is assumed to be an Itô semimartingale process that can be written as:

$$X_t = X_0 + \int_0^t b_s ds + \int_0^t \sigma_s dB_s + \sum_{s\leq t} \Delta X_s, \tag{1}$$

where $X_0 + \int_0^t \mu_s ds + \int_0^t \sigma_s dB_s$ is the continuous semimartingale component of the process, which is the sum of a local martingale plus an adapted process with finite variation component. Additionally, ΔX_s is a jump at time s, defined as:

$$\Delta X_s = X_s - \lim_{\tau<s, \tau\to s} X_\tau.$$

Given this definition, the jump part of X_t in the time interval $[0, t]$ is defined to be $\sum_{s\leq t} \Delta X_s$. Note that when the jump is a Compound Poisson Process (CPP) - i.e. a finite activity jump process - then it can be expressed as:

$$J_t = \sum_{s\leq t} \Delta X_s = \sum_{i=1}^{N_t} Y_i,$$

where N_t is number of jumps in $[0, t]$, N_t follows a Poisson process, and the Y_i's are i.i.d. and are the sizes of the jumps. The CCP assumption has been widely used in the literature on modeling, forecasting, and testing for jumps. However, as discussed above, recent evidence suggests that processes may contain infinite activity jumps - i.e. infinite tiny jumps that look similar to continuous movements. In such cases, the CCP assumption is clearly violated, and hence we draw in such cases on the theory of Jacod (2008) and Aït-Sahalia and Jacod (2009b,c) when applying standard BNS (2006) type jump tests in the sequel. The empirical evidence discussed in this paper involves examining the structure of the jump component of the log return process, X_t, using one historically observed price sample path $\{X_0, X_{\Delta_n}, X_{2\Delta_n} ... X_{n\Delta_n}\}$, where Δ_n is deterministic. The increment of the process at time $i\Delta_n$ is denoted by:

$$\Delta_i^n X = X_{i\Delta_n} - X_{(i-1)\Delta_n}.$$

For convenience, we consider the case $t = n\Delta_n$ in the sequel.[21]. In general, integrated volatility and quadratic variation are formally defined as:

$$IV_t = \int_0^t \sigma_s^2 ds = [\text{variation due to continuous component}]_t$$

$$QV_t = [X, X]_t = \int_0^t \sigma_s^2 ds + \sum_{t-1 \leq s < t} \Delta X_s^2 = IV_t + [\text{variation due to jump component}]_t$$

Volatility processes QV_t and IV_t are not observable. However, one could exploit high or ultra high frequency data in financial markets to estimate these variables. If the process is continuous, $IV_t = QV_t$ and their noisy estimators, hereby referred to as realized measures (RM) could be written as

$$RM_{t,n} = IV_t + N_{t,n}$$

where $N_{t,n}$ denotes the measurement error associated with the realized measure $RM_{t,n}$. There are two sources of measurement errors. One is due to the so-called microstructure noise effect of high frequency data, and the second is due to standard noise. There are a few realized measures and methods proposed to alleviate the effect of the contaminated high frequency noise. For example, Corradi, Distaso and Swanson (2009, 2010) derive consistent estimators of predictive conditional densities of integrated volatility using these noisy realized measures. They show that by choosing an appropriate realized measure, one can achieve consistent estimation, even in the presence of jumps and microstructure noise in prices. They thereby construct conditional predictive densities and confidence intervals for integrated volatility using realized measures, which may be of interest to volatility derivatives traders. Note that as microstructure noise is not the focus of our paper, we will focus mainly on three key realized measures that are commonly used, i.e. realized volatility, bipower, and tripower variation (many papers in the extant literature now look also at multipower variation). For a list of other realized measures that are robust to microstructure noise, see for instance Corradi, Distaso and Swanson (2009, 2010). The realized volatility of equity return process X is defined as follows:

$$RV_{t,n} = \sum_{i=1}^{n} (\Delta_i^n X)^2$$

It has been shown that when n is large, realized volatility converges almost surely to the

[21] See Jacod (2008) for further details.

quadratic variation of the process[22]. If we measure volatility within a day then,

$$RV_{t,n} = \sum_{i=1}^{n} (\Delta_i^n X)^2 \simeq \int_{t-1}^{t} \sigma_s^2 ds + [\text{variation due to jump component between } day \text{ day } t-1 \text{ and } t]_t$$

were n here is used to denote the number of incremental returns within a day or any other fixed time period. Multipower variations are constructed on the basis of

$$V_{r_1,r_2...,r_j} = \sum_{i=j+1}^{n} |\Delta_i^n X|^{r_1} |\Delta_{i-1}^n X|^{r_2} .... |\Delta_{i-j}^n X|^{r_j}.$$

where $r_1, r_2, ..., r_j$ are positive, such that $\sum_1^j r_i = k$. Bipower (BV) and tripower variation (TP) are special cases of multipower variations. Specifically, Bipower variation is given by

$$BV_{t,n} = (\mu_1)^{-1} \sum_{i=2}^{n} V_{1,1} \simeq \int_{t-1}^{t} \sigma_s^2 ds$$

where $\mu_1 = E|Z| = 2^{1/2}\Gamma(1)/\Gamma(1/2)$ and Z is a standard normal random variable; and tripower variation is given by

$$TV_{t,n} = V_{\frac{2}{3},\frac{2}{3},\frac{2}{3}} \mu_{\frac{2}{3}}^{-3} \simeq \int_{t-1}^{t} \sigma_s^2 ds$$

where $\mu_r = E(|Z|^r)$and Z is a $N(0,1)$ random variable. Finally, to illustrate the nuts and bolts of microstructure robust realized measures, we include the formula for a commonly used subsample based realized volatility measure, $\widehat{RV}_{t,l,M}$, suggested by Zhang, Mykland and Ait-Sahalia (2005), and defined as

$$\widehat{RV}_{t,l,n} = RV_{t,l,n}^{avg} - 2l\widehat{v}_{t,n},$$

where

$$\widehat{v}_{t,n} = \frac{RV_{t,l,n}}{2n} = \frac{1}{2n} \sum_{j=1}^{n} (\Delta_i^n X)^2$$

and

$$RV_{t,l,n}^{avg} = \frac{1}{B} \sum_{b=1}^{B} RV_{t,l}^{b} = \frac{1}{B} \sum_{b=1}^{B} \sum_{j=1}^{t-1} (X_{t+\frac{jB+b}{n}} - X_{t+\frac{(j-1)B+b}{n}})^2$$

Here, $Bl \cong n, l = O(n^{1/3})$

Forecasting Realized Volatility With the availability of high frequency data in recent years, much effort has been dedicated to building good models to forecast realized volatility, a "model free" estimator of ex post variance. To see the link of this type of forecast to the volatility forecasting

[22] See Barndorff-Nielsen and Shephard (2002).

in discrete time models, note that the daily volatility implied by a discrete time model, let's say GARCH, is equivalent to $\sqrt{QV_t}$ or $\sqrt{IV_t}$ (without jumps) in continuous time[23], which are proxied by$\sqrt{RV_{t,n}}$ or $\sqrt{BV_{t,n}}$ or $\sqrt{TV_{t,n}}$. The link justifies the rationale of this type of forecast in the literature. We highlight several important papers in this area of research. In a key paper is Andersen, Bollerslev, Diebold, Laby (2003), who show empirically that a long memory Gaussian VAR can capture the dynamics of volatility. They apply a simple trivariate VAR (VAR-RV) to model the dynamics of volatilities of logarithmic exchange rates, i.e. DM/$, Yen/$ and Yen/DM. In particular, denote y_t as the vector of the the exchange rates, and define the forecasting equation as

$$\Phi(L)(1-L)^d(y_t-\mu)=\varepsilon_t$$

where ε_t is a vector white noise process. The authors point out that this simple framework performs better than many alternative models that have been used in this literature. In the same regression model, Corsi (2003) proposes the so called HAR - RV model in which realized volatility is specified as a linear sum of the lagged realized volatilities over different horizons, i.e.

$$RV_{t,t+h}=\beta_0+\beta_d RV_t+\beta_w RV_{t,t-5}+\beta_m RV_{t,t-22}+\epsilon_{t+h}$$

where h is the forecasting horizon, i.e. $h=1,5,22.$ and $RV_{t,t+h}=h^{-1}[RV_{t+1}+RV_{t+2}+...+RV_{t+h}]$. ABD (2007) generalizes HAR - RV to linear and nonlinear HAR-RV, HAR-RV-J and HAR-RV-CJ. In particular, the new feature of the model that they propose involves incorporating the variation of jump components of the log price process into their forecasts. The advantage of these models is that they are rather simple to implement via least squares estimation and they take advantage of recent developments in the construction of robust jump tests. In the next section, we summarize robust jump testing and discuss the quantification of large jumps and small jumps, a departure that can potentially help improve the model's forecasting performance.

[23] To see this, in the case of continuous process, $Var(X_t-X_{t-1})=E(\int_{t-1}^{t}\sigma_s dB_s)^2=\int_{t-1}^{t}\sigma_s^2 ds$. In addition, for convenience in notation, in this section, $\sqrt{QV_t}$ or $\sqrt{IV_t}$ are refered to as volatilities in a day.

.

4 Volatility and Jumps

Thus far, we have summarized important developments in volatility models, with focus on time varying and stochastic volatility as well as nonparametric volatility estimators. All of our models share the same feature that volatilities are unobserved, or belong to class of missing variables. We now turn to a discussion of jumps, testing for jumps, and disentangling the effects of jumps from measures of volatility. This section also contains the results of our empirical analysis of jumps and volatility. Much of the discussion in this section is taken from Duong and Swanson (2010).

4.1 Testing for Jumps

In this section, we review some theoretical results relating to testing for jumps, namely testing whether $J_t = \sum_{s \leq t} \Delta X_s \neq 0$. In pioneering work, BNS (2006) proposes a robust and simple test for a class of Brownian Itô Semimartingales plus Compound Poisson jumps. In recent work, Aït-Sahalia and Jacod (2009b) among others develop a different test which applies to a large class of Itô-semimartingales, and allows the log price process to contain infinite activity jumps - small jumps with infinite concentrations around 0. In this paper, we follow the jump test methodology of Huang and Tauchen (2005) as well as Barndorff-Nielsen and Shephard (2006), which looks at the difference between the continuous component and total quadratic variation in order to test for jumps. However, we make use of the limit theorems developed and used by Jacod (2008) and Aït-Sahalia and Jacod (2009b) in order to implement the Barndorff -Nielsen and Shephard (2006) type test under the presence of both infinite activity and finite activity jumps (see Section 4 for further discussion). A simplified version of the results of the above authors applied to (1) for the one-dimensional case is as follows. If the process X is continuous, let $f(x) = x^n$ (exponential growth), let ρ_{σ_s} be the law $N(0, \sigma_s^2)$, and let $\rho_{\sigma_s}(f)$ be the integral of f with respect to this law. Then:

$$\sqrt{\frac{1}{\Delta_n}} \left(\Delta_n \sum_{i=1}^{n} f(\frac{\Delta_i^n X}{\sqrt{\Delta_n}})^2 - \int_0^t \rho_{\sigma_s}(f) ds \right) \xrightarrow{L-S} \int_0^t \sqrt{\rho_{\sigma_s}(f^2) - \rho_{\sigma_s}^2(f)} dB_s \tag{2}$$

Here, $L - S$ denotes stable convergence in law, which also implies convergence in distribution. For $n = 2$, the above result is the same as BNS (2006). More generally:

$$\sqrt{\frac{1}{\Delta_n}} \left(\sum_{i=1}^{n} (\Delta_i^n X)^2 - \int_0^t \sigma_s^2 ds \right) \xrightarrow{D} N(0, \int_0^t \vartheta \sigma_s^4 ds) \tag{3}$$

or

$$\frac{\sqrt{\frac{1}{\Delta_n}}\left(\sum_{i=1}^{n}(\Delta_i^n X)^2 - \int_0^t \sigma_s^2 ds\right)}{\sqrt{\int_0^t \vartheta \sigma_s^4 ds}} \xrightarrow{D} N(0,1), \tag{4}$$

where ϑ is constant and where $\int_0^t \sigma_s^2 dt$ is known as the integrated volatility or the variation of the continuous component of the model and $\int_0^t \sigma_s^2 dt$ is integrated quarticity. From the above result, notice that if the process does not have jumps, then $\sum_{i=1}^{n}(\Delta_i^n X)^2$,which is an approximation of quadratic variation of the process, should be "close" to the integrated volatility. This is the key idea underlying the BNS (2006) jump test. For appropriate central limit theorems, in tests with both finite and infinite activity jumps, refer to Barndorff-Nielsen, Graverson, Jacod, Podolskij, and Shephard (2006), in the case of continuous semimartingales and Barndorff-Nielsen, Shephard, and Winkel (2006) for discontinuous process wih Lévy jumps. A final crucial issue in this jump test is the estimation of $\int_0^t \sigma_s^2 dt$ and $\int_0^t \sigma_s^4 dt$ in the presence of both finite and infinite activity jumps. As remarked in BNS (2006), in order to ensure that tests have power under the alternative, integrated volatility and integrated quarticity estimators should be consistent under the presence of jumps. The authors note that robustness to jumps is straightforward so long as there are a finite number of jumps, or in cases where the jump component model is a Lévy or non-Gaussian OU model (Barndorff-Nielsen, Shephard, and Winkel (2006)). Moreover, under infinite activity jumps, note that as pointed out in Jacod (2007), there are available limit results for volatility and quarticity estimators for the case of semimartingales with jumps.

Turning again to our discussion of volatility and quarticity, note that in a continuation of work initiated by Barndorff-Nielsen and Shephard (2004), Barndorff-Nielsen, Graverson, Jacod, Podolskij, and Shephard (2006) and Jacod (2007) develop general so-called multipower variation estimators of $\int_0^t \sigma_s^k ds$, in the case of continuous semimartingales and semimartingales with jumps, respectively, which are based on

$$V_{r_1,r_2...,r_j} = \sum_{i=j+1}^{n} |\Delta_i^n X|^{r_1} |\Delta_{i-1}^n X|^{r_2} .... |\Delta_{i-j}^n X|^{r_j}.$$

where $r_1,r_2,...,r_j$ are positive, such that $\sum_1^j r_i = k$. For cases where $k = 2$ and $k = 4$, BNS (2006) use $V_{1,1}$ (bipower variation) and $V_{1,1,1,1}$. In our jump test implementation, we use $V_{\frac{2}{3},\frac{2}{3},\frac{2}{3}}$(tripower variation) and $V_{\frac{4}{5},\frac{4}{5},\frac{4}{5}}$. The reason we use tripower variation, $V_{\frac{2}{3},\frac{2}{3},\frac{2}{3}}$, instead of bipower variation, $V_{1,1}$, is that it is more robust to clustered jumps. Denote the estimators of $\int_0^t \sigma_s^2 ds$ and $\int_0^t \sigma_s^4 ds$ to

be $\widehat{IV}$ and $\widehat{IQ}$, and note that:

$$\widehat{IV} = V_{\frac{2}{3},\frac{2}{3},\frac{2}{3}}\mu_{\frac{2}{3}}^{-3} \simeq \int_0^t \sigma_s^2 ds \tag{5}$$

and

$$\widehat{IQ} = \Delta_n^{-1} V_{\frac{4}{3},\frac{4}{3},\frac{4}{3}}\mu_{\frac{4}{5}}^{-5} \simeq \int_0^t \sigma_s^4 ds, \tag{6}$$

where $\mu_r = E(|Z|^r)$and Z is a $N(0,1)$ random variable.

Regardless of the estimator that is used, the appropriate test hypotheses are:

$$H_0 : X_t \text{ is a continuous process}$$

$$H_1 : \text{the negation of } H_0 \text{ (there are jumps)}$$

If we use multi-power variation, under the null hypothesis the test statistic which directly follows from the CLT mentioned above is:

$$LS_{jump} = \frac{\sqrt{\frac{1}{\Delta_n}}\left(\sum_{i=1}^n (\Delta_i^n X)^2 - \widehat{IV}\right)}{\sqrt{\vartheta \widehat{IQ}}} \xrightarrow{D} N(0,1)$$

and the so-called jump ratio test statistic is:

$$RS_{jump} = \frac{\sqrt{\frac{1}{\Delta_n}}}{\sqrt{\vartheta \widehat{IQ}/(\widehat{IV})^2}}\left(1 - \frac{\widehat{IV}}{\sum_{i=1}^n (\Delta_i^n X)^2}\right) \xrightarrow{D} N(0,1).$$

Of note is that an adjusted jump ratio statistic has been shown by extensive Monte Carlo experimentation in Huang and Tauchen (2005), in the case of CCP jumps, to perform better than the two above statistics, being more robust to jump over-detection. This adjusted jump ratio statistic is:

$$AJ_{jump} = \frac{\sqrt{\frac{1}{\Delta_n}}}{\sqrt{\vartheta max(t^{-1}, \widehat{IQ}/(\widehat{IV})^2)}}\left(1 - \frac{\widehat{IV}}{\sum_{i=1}^n (\Delta_i^n X)^2}\right) \xrightarrow{L} N(0,1) \tag{7}$$

In general if we denote the daily test statistics to be $Z_{t,n}(\alpha)$,where n is the number of observations per day and α is the test significance level [24], then we reject the null hypothesis if $Z_{t,n}(\alpha)$ is in excess of the critical value Φ_α, leading to a conclusion that there are jumps. The converse holds if $Z_{t,n}(\alpha)$ is less than Φ_α. In our empirical application, $Z_{t,n}(\alpha)$ is the adjusted jump ratio statistic,

[24] i.e., $\Delta_n = 1/n$

and we calculate the percentage of days that have jumps, for the period from 1993 to 2008. We now turn to a discussion of large jumps and constructing measures of the daily variation due to continuous and jump components.

4.2 Large Jumps and Small Jumps

There is now clear evidence that jumps are prevalent in equity market. For example, Huang and Tauchen (2005) construct the above jump test statistics, and find that jumps contribute about 7% to the total variation of daily stock returns. Aït-Sahalia and Jacod (2009b) not only find jumps but given the existence of jumps, they look more deeply into the structure of the jumps, and for Intel and Microsoft returns they find evidence of the existence of infinite activity jumps.

An important focus in our paper is to the decomposition of jumps into "large" and "small" components so that we may assess their contributions to the overall variation of the price process. In particular, for some fixed level γ, define large and small jump components as follows, respectively:

$$LJ_t(\gamma) = \sum_{s \leq t} \Delta X_s I_{|\Delta X_s| \geq \gamma}.$$

and

$$SJ_t(\gamma) = \sum_{s \leq t} \Delta X_s I_{|\Delta X_s| < \gamma}.$$

The choice of γ may be data driven, but in this paper we are more concerned with scenarios where there is some prior knowledge concerning the magnitude of γ. For example, under various regulatory settings, capital reserving and allocation decisions may be based to a large extent on the probability of jumps or shocks occurring that are of a magnitude greater than some known value, γ. In such cases, planners may be interested not only in knowledge of jumps of magnitude greater than γ, but also in characterizing the nature of the variation associated with such large jumps. The procedure discussed in this section can readily be applied to uncover this sort of information.

4.3 Realized measures of daily jump variation

The partitioning of variation due to continuous and jump components can be done, for example, using truncation based estimators which have been developed by Mancini (2001,2004,2009) and

Jacod (2008). One can also simply split quadratic variation into continuous and jump components by combining various measures of integrated volatility, such as bipower or tripower variation and realized volatility. Andersen, Bollerslev, and Diebold (2007) do this, and construct measures of the variation of the daily jump component as well as the continuous component. In this paper we use their method, but apply it to both small and large jumps. In particular, once jumps are detected, the following risk measures introduced by Andersen et al. (2007) are constructed:

VJ_t = *Variation of the jump component* $= max\{0, RV_t - \widehat{IV_t}\} * I_{jump,t}$

VC_t = *Variation of continuous component* $= RV_t - VJ_t$,

where $RV_t = \sum_{i=1}^{n}(\Delta_i^n X)^2$ is the daily realized volatility (i.e. a measure of the variation of the entire (log) stock return process), I_{jump} is an indicator taking the value 0 if there are no jumps and 1 otherwise, and n is the number of intra-daily observations. One can then calculate daily jump risk. Note that in these formulae, the variation of the continuous component has been adjusted (i.e. the variation of the continuous component equals realized volatility if there are no jumps and equals $\widehat{IV_t}$ if there are jumps). In addition, note that $\sum_{i=1}^{n}(\Delta_i^n X)^2 I_{|\Delta_i^n X|\geq\gamma}$converges uniformly in probability to $\sum_{s\leq t}(\Delta X_s)^2 I_{|\Delta X_s|\geq\gamma}$,as n goes to infinity [25] Thus, the contribution of the variation of jumps with magnitude larger than γ and smaller than γ are denoted and calculated as follows:

Realized measure of large jump variation: $VLJ_{t,\gamma}$=min$\{VJ_t, (\sum_{i=1}^{n}(\Delta_i^n X)^2 I_{|\Delta_i^n X|\geq\gamma} * I_{jump,t})\}$,

Realized measure of small jump variation: $VSJ_{t,}\gamma = VJ_t - VLJ_{t,\gamma}$,

where I_{jump} is defined above and $I_{jump,\gamma}$ is an indicator taking the value 1 if there are large jumps and 0 otherwise. This condition simply implies that large jump risk is positive if the process has jumps and has jumps with magnitude greater than γ.

Now we can write the relative contribution of the variation of the different jump components to total variation in a variety of ways:

Relative contribution of continuous component $= \frac{VC_t}{RV_t}$

Relative contribution of jump component $= \frac{VJ_t}{RV_t}$

Relative contribution of large jump component $= \frac{VLJ_{t,\gamma}}{RV_t}$

Relative contribution of small jump component $= \frac{VLS_{t,\gamma}}{RV_t}$

Relative contribution of large jumps to jump variation $= \frac{VLJ_{t,\gamma}}{VJ_t}$

Relative contribution of small jumps to jump variation $= \frac{VLS_{t,\gamma}}{VJ_t}$

[25] See Jacod (2008), Aït-Sahalia and Jacod (2009a) for further details.

We are now ready to discuss some empirical findings based on the application of the tools discussed in this section.

5 Empirical Findings

5.1 Data description

We use a large tick by tick data set of 25 DOW 30 stocks available for the period 1993-2008. The data source is the TAQ database. We use only 25 stocks because we purge our data set of those stocks that not frequently traded or are not available across the entire sample period. For the market index, we follow several other papers and look at S&P futures. We also follow the common practice in the literature of eliminating from the sample those days with infrequent trades (less than 60 transactions at our 5 minute frequency).

One problem in data handling involves determining the method to filter out an evenly-spaced sample. In the literature, two methods are often applied - *previous tick* filtering and *interpolation* (Dacorogna, Gencay, Müller, Olsen, and Pictet (2001)). As shown in Hansen and Lund (2006), in applications using quadratic variation, the *interpolation* method should not be used, as it leads to realized volatility with value 0 (see *Lemma 3* in their paper). Therefore, we use the *previous tick* method (i.e. choosing the last price observed during any interval). We restrict our dataset to regular time (i.e. 9:30am to 4:00pm) and ignore ad hoc transactions outside of this time interval. To reduce microstructure effects, the suggested sampling frequency in the literature is from 5 minutes to 30 minutes[26]. As mentioned above, we choose the 5 minute frequency, yielding a maximum of 78 observations per day.

5.2 Jump and Large Jump Results

We implement our analysis in two stages. In the first stage we test for jumps and in second stage we examine large jump properties, in cases where evidence of jumps is found. The list of the companies for which we examine asset returns is given in Table 1, along with a summary of our jump test findings. The rest of the tables and figures summarize the results of our empirical investigation. Before discussing our findings, however, we briefly provide some details about the calculations that

[26] See Aït-Sahalia, Y., Mykland, P. A., and Zhang, L. (2005)

we have carried out.

All daily statistics are calculated using the formulae in Section 4 with:

$$\Delta_n = \frac{1}{n} = \frac{1}{\#\text{ of 5 minute transactions / day}}$$

Therefore, $\Delta_n = 1/78$ for most of the stocks in the sample, except during various shortened and otherwise nonstandard days, and except for some infrequently traded stocks. This also implies the choice of time to be the interval $[0, 1]$, where the time from $[0, 1]$ represents the standardizing time with beginning (9 am) set to 0 and end (4.30 pm) set to 1. In our calculations of estimates of integrated volatility and integrated quarticity, we use multipower variation, as given in (5) and (6). Recall also that $\Delta_i^n X = X_{i\Delta_n} - X_{(i-1)\Delta_n}$ is simply the incremental return of $X_{i\Delta_n}$. For any trading day, X_0 and X_1 correspond to the first and the last observations of the day. Denote T as the number of days in the sample. We construct the time series $\{Z_{t,n}(\alpha)\}_{t=0}^{T}$ and $\left\{\frac{VC_t}{RV_t}, \frac{VJ_t}{RV_t}, \frac{VLJ_{t,\gamma}}{RV_t}, \frac{VSJ_{t,\gamma}}{RV_t}\right\}_{t=0}^{T}$. The number of days and proportion of days identified as containing jumps can easily be calculated as:

Number of days identified as jumps $= \sum_{i=0}^{T} I(Z_{i,n}(\alpha) > \Phi_\alpha)$.

Proportion of days identified as jumps $= \frac{\sum_{i=0}^{T} I(Z_{t,n}(\alpha) > \Phi_\alpha)}{T}$.

In addition, we construct the following monthly time series

Proportion of days identified as jumps in a month$= \frac{\sum_{i=m}^{m+h} I(Z_{t,n}(\alpha) > \Phi_\alpha)}{h}$

Monthly average relative contribution of jump component $= \frac{\sum_{i=m}^{m+h} \frac{VJi}{RVi}}{h}$

Monthly average relative contribution of large jump component truncated at level $\gamma = \frac{\sum_{i=m}^{m+h} \frac{VLJ_{i,\gamma}}{RV_t}}{h}$,

where m is the starting date and h is the number of days in each month. On average, there are 22 business days per month. Note that there are 12 statistics each year for each time series.

Here, $I(\cdot)$ denotes the indicator function. The average relative contribution of continuous, jump, and large jump components to the variation of the process is reported using the mean of the sample (i.e. we report the means of $\frac{VC_t}{RV_t}, \frac{VJ_t}{RV_t}, \frac{VLJ_{t,\gamma}}{RV_t}$, and $\frac{VSJ_{t,\gamma}}{RV_t}$).

In addition to reporting findings based on examination of the entire sample period, we also split the sample into two periods. The first period is from 1993 to 2000 and the second period is from 2001 to 2008. The reason for doing this is that we would like to see whether jump activity changes over time. Moreover, these subsamples correspond roughly to break dates for financial data found in Cai and Swanson (2010).

In the sequel, we provide figures for representative individual stocks in our sample (i.e. Walmart, IBM, Bank of America and Citigroup). These stocks are chosen on the basis of their market systematic risk beta. Namely, Walmart has low beta of around 0.3, IBM has a beta close to 0.7, and Bank of America and Citigroup are more risky stocks with betas of around 2.6 and 2.8.

Turning now to our results, a first sense of the prevalence of jumps can be formulated by inspecting Panels A,B, C and D of Figure 1, where statistics higher than 3.09 (i.e. the 0.001 significance level critical value) are presented for the entire sample from 1993 to 2008. It is obvious that jumps are prevalent. Additionally, it should be noted that there is a marked difference in jump frequency between 1993-2000 and 2001-2008, where the first period is much more densely populated with jumps than the latter period. The highest statistic values are around 11, for Walmart in 1997, 11 for IBM in 1994, 10 for Bank of America in 1996 and 7 for Citigroup from 1996 to 1998. Post 2000, the highest statistics are consistently located in 2002 and 2006-2008. Moreover, a simple visual check of the statistic magnitudes in this figure suggests that jumps are more prevalent in the earlier sample period, with respect to both frequency and significance level (more will be said on this later).

Regarding our choice of the large jumps, an important step is to choose truncation levels, γ. If we choose arbitrarily large truncation levels, then clearly we will not find evidence of large jumps. Also one may easily proceed by just picking the truncation level based on the percentiles of the entire historical sample of the 5 minute log return. However, results could then turn out to be difficult to interpret, as in one case the usual choice of $90th$ or $75th$ percentiles leads to virtually no large jumps while the choice $25th$ or $10th$ percentiles leads to a very large number of large jumps. In addition, "large" jumps are often thought of as abnormal events that arise at a frequency of one in several months or even years. Therefore, a reasonable way to proceed is to pick the truncation level on the basis of the sample of the monthly maximum increments - monthly based abnormal events. Specifically, we set three levels $\gamma = 1, 2, 3$ to be the $50th, 75th$, and $90th$ percentiles of the entire sample from 1993 to 2008. Panels A,B,C, and D of Figure 2 depict the monthly largest absolute increments and the jump truncation levels used in our calculations of the variation of large and small jump components. Again, it is quite obvious that the monthly maximum increments are dominant in the previous decade. The larger monthly increments in current decades are mostly located in 2006-2008 and 2002-2004. As a result, the fixed truncation levels which are chosen across the entire sample result in more "hits" in previous decade than in the current one. The truncation

level of Citigroup is the largest of the four stocks depicted (for example at $\gamma = 3$ the level is approximately 0.04 for Citigroup and 0.025 for IBM).

Notice that the graphs in Figures 3A and 3B depict magnitudes of the variation of continuous, jump, and truncated jump components of returns for our 4 sample stocks. Namely, the plots are of daily realized volatility, and realized variance of continuous, jump and large jump components at different truncation levels. As might be expected, inspection of the graphs suggests a close linkage between the greater number of jumps in the first decade of the sample and the and large jump risk over the same period. For example, in the case of IBM, the variation of the jump components is clearly dominant in the earlier decade. The highest daily jump risk occurs in late 1998, and is above 0.018. Indeed, at jump truncation level 3, we only see large jump risk for the years 1994, 1996, 1998, 2000 and 2008. Combined with the results of Figure 1, this again strongly suggests that there was much more turbulence in the earlier decade.

Turning now to our tabulated results, first recall that Table 1 reports the proportion of days identified as having jumps, at 6 different significant levels, $\alpha = \{0.1, 0.05, 0.01, 0.005, 0.001, 0.0001\}$. Again, there is clear evidence of jumps in both periods. However, the jump frequency in the 1993-2000 sample is significantly higher than that in the 2001-2008 sample, across all stocks and test significance levels. For example, at the $\alpha = 0.005$ and 0.001 levels, the average daily jump frequencies are 46.9% and 22.8% during the 1993-2000 period, as compared with 16.8% and 9.4% during the 2001-2008 period, respectively. When considering individual stocks, the story is much the same. As illustrated in Figures 1, and tabulated in Table 1, the proportion of "jump-days" for IBM and for the Bank of America are 5.9% and 8.8% during the 2000s, which is much smaller than the value of 19.2% and 21.3% for the two stocks during the 1990s, based on tests implemented using a significance level of $\alpha = 0.001$.

Of course, when calculating jump frequencies, we ignore the magnitudes of the jumps. Table 2 addresses this issue by summarizing another measure of jumps - namely the average percentage contribution of jumps to daily realized variance. Details of the measures reported are given above and in Section 4. In support of our earlier findings, it turns out that jumps account for about 15.6% and 8.1% of total variation at significance levels $\alpha = 0.005$ and 0.001, respectively, when considering the entire sample period from 1993-2008. Moreover, analogous statistics for the period 1993-2000 are 25.1% and 12.7%, while those for the 2001-2008 period are 7% and 5%. The statistics for IBM and Bank of America are 25.3% and 10.7% for the period 1993-2000 and 3.5% and 2.3% for the

period 2001-2008 while those for the entire samples are 7.9% and 6.6%. This result is consistent with our earlier findings through figure analysis.

In summary, without examining the impact of large jumps, we already have evidence that: (i) There is clear evidence that jumps characterize the structure of the returns of all of the stocks that we examine. (ii) The 1990s are characterized by the occurrence of more jumps than the 2000s. (iii) The contribution of jumps to daily realized variance is substantively higher during the 1990s than the 2000s. (iv) Our results are consistent across all stocks, suggesting the importance of jump risk comovement during turbulence periods.

In our empirical analysis of large jumps, we carry out the same steps as those employed above when examining overall jump activity. Results are reported in Tables 3A-C are for truncation levels $\gamma = 1, 2, 3$ at 6 different significant levels, $\alpha = \{0.1, 0.05, 0.01, 0.005, 0.001, 0.0001\}$. As mentioned earlier, Figures 1 and 3 contain plots of jump test statistics and realized variation not only for overall jump activity, as discussed above, but also for large jumps. Examination of these tables suggest a number of conclusions.

Across the entire sample, there is evidence of large jumps at all levels by measure of variation. Table 3A reports the proportion of days identified as having large jumps for truncation level $\gamma = 1$. It can be seen that the proportion of variation due to large jumps at truncation level $\gamma = 1$ accounts for about 0.9% and 0.6% of total variation (regardless of stock), at significance levels $\alpha = 0.005$ and 0.001, respectively. Values at significance level 0.001 for the periods 1993-2000 and 2001-2008 are around 0.8% and 0.4%, respectively. For $\gamma = 2$, values are 0.4% and 0.3% at significant levels $\alpha = 0.005$ and 0.001, respectively, when considering the entire sample. Values at significance level 0.001 for the periods 1993-2000 and 2001-2008 are around 0.4% and 0.2% for period 1993-2000 and 2001-2008, respectively. A similar result obtains for $\gamma = 3$, suggesting that large jump variability is around twice as big (as a proportion of total variability) for the latter sub-sample, regardless of truncation level. As previously, these results are surprisingly stable across stocks. Although not included here, our analysis of the market index data discussed above yielded a similar result. Further examination of the statistics in the Tables 3A-C also yields another interesting finding. In particular, though proportions of jumps and large jumps at truncation level $\gamma = 1, 2, 3$ are all larger in the previous decade, the difference is smaller and increasingly narrows as higher truncation levels are considered, when examining large jumps. This result, which is true for many of our stocks, suggests an increased role of large jumps in explaining daily realized variance during the

latter sub-sample. To illustrate this point, which is apparent upon inspection of average statistics constructed for all 25 stocks, we investigate the case of of ExxonMobil, where we look at all statistics at significance level $\alpha = 0.001$. The proportion of variation of jumps to total variation is 17% for the period 1993-2000 (as shown in Table 2), almost 3 times as much as the corresponding value of 6.2% in 2001-2008. However for large jumps at truncation level $\gamma = 1$, the analogous value is 0.6% for 1993-2008, which is just 1.5 times as much as the 0.4% value during 2001-2008. Similarly at truncation level $\gamma = 2$, the value is 0.4% for 1993-2008 and 0.2% for 2001-2008. Interestingly, at truncation level $\gamma = 3$, the proportion of variation of jumps is 0 for period 1993-2000 while it is 0.1% for period 2001-2008. Therefore, with respect to large jump we find that: (i) Large jumps incidence and magnitudes are consistent with our earlier finding that the 1990s are much more turbulent than the 2000s. (ii) However, for higher truncation levels, the contribution of jump risk during the two periods becomes much closer, and indeed the contribution during the latter period can actually become marginally greater. This suggests that while the overall role of jumps is lessening, the role of large jumps has not decreased, and indeed, the relative role of large jumps, as a proportion of overall jumps has actually increased in the 2000s.

6 Concluding Remarks

In this paper we review some of the recent literature on volatility modelling and jumps, with emphasis on the notion that these variables are unobserved latent variables, and thus can be viewed in some sense as "missing data". Many estimators of volatility, both continuous and discrete, as well as both parametric and nonparameteric are also reviewed..

In an empirical investigation, we provide new evidence of jumps in individual log price processes, and note that there are clearly comovements during turbulent times, for all stocks. More noticeably, jump incidence is greater during the 1990s than during the 2000s, although the incidence of "very large" jumps is similar across both decades, and the relative importance of large jumps has increased.

References

Aït-Sahalia, Y. (1996). Testing Continuous-Time Models of the Spot Interest Rate. *Review of Financial Studies* 9(2), 385-426.

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. and J. Jacod (2009a). Analyzing the Spectrum of Asset Returns: Jump and Volatility Components in High Frequency Data. Working Paper, Princeton University.

Aït-Sahalia, Y. and J. Jacod (2009b). Testing for Jumps in a Discretely Observed Process. *Annals of Statistics* 37, 184-222.

Aït-Sahalia, Y. and J. Jacod (2009c). Estimating the Degree of Activity of Jumps in High Frequency Data. *Annals of Statistics* 37, 2202-2244.

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

Andersen, T.G. and T. Bollerslev (1998a). Deutsche Mark-Dollar Volatility: Intraday Activity Patterns, Macroeconomic Announcements, and Longer Run Dependencies. *Journal of Finance* 53, 219-265.

Andersen, T.G., T. Bollerslev, and F.X. Diebold (2007). Roughing it Up: Including Jump Components in the Measurement, Modeling and Forecasting of Return Volatility. *Review of Economics and Statistics* 89, 701-720.

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

Andersen, T.G. and J. Lund (1997). Estimating Continuous-Time Stochastic Volatility Models of the Short Term Interest Rate Diffusion. *Journal of Econometrics* 77, 343-377.

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., S.E. Graversen, J. Jacod, M. Podolskij, and N. Shephard (2006). A Central Limit Theorem for Realized Power and Bipower Variations of Continuous Semimartingales, in: **From Stochastic Analysis to Mathematical Finance**, Festschrift for Albert Shiryaev, Springer and Verlag, Y. Kabanovand R. Lipster (eds.): New York: U.S.A.

Barndorff-Nielsen, O.E. and N. Shephard (2002). Econometric Analysis of Realized Volatility and Its Use in Estimating Stochastic Volatility Models. *Journal of Royal Statistical Society*, 64,

253–80.

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

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 (2006). Limit Theorems for Multipower Variation in the Presence of Jumps. *Stochastic Processes and Their Applications* 116, 796-806.

Barone-Adesi, G., R.F. Engle, and L. Mancini (2008). A GARCH Option Pricing Model with Filtered Historical Simulation. *Review of Financial Studies* 21, 1223–1258.

Bates, D.S. (1996). Jump and Stochastic Volatility: Exchange Rate Processes Implicit in Deutsche Mark Options. *Review of Financial Studies* 9, 69-107.

Bates, D.S. (2000). Post-87 Crash Fears in the S&P 500 Futures Option Market. *Journal of Econometrics* 94, 181-238

Bates, D.S. (2003). Empirical Option Pricing: a Retrospection. *Journal of Econometrics* 116, 387–404.

Beckers, S. (1980). The Constant Elasticity of Variance Model and US Implications for Option Pricing. *Journal of Finance* 35, 661-673.

Billingsley, P. (1968). **Convergence of Probability Measures**. New York: Wiley.

Black, F. (1976). Studies in Stock Prie Volatility Changes. *In Proceedings of the 1976 Meetings of the Business and Economic Statistics Section*, 177-181. American Statistical Association.

Black, F. and M. Scholes (1973). The Pricing of Options and Corporate Liabilities. *Journal of Political Economy* 81, 637-654.

Bollerslev, T. (1986). Generalized Autogressive Conditional Heteroskedasticity. *Journal of Econometrics* 31, 307-327.

Bollerslev, T., R. Chou, and K. Kroner (1992). ARCH Modeling in Finance: A Review of Theory and Empirical Evidence. *Journal of Econometrics* 52, 5-59.

Bollerslev, T., R.F. Engle, and D.B. Nelson (1994). ARCH Models. In R. Engle and D. McFadden (eds.), **Handbook of Econometrics**, vol. 4. Amsterdam: North-Holland.

Bollerslev, T., R.F. Engle, and J.M. Wooldridge (1988). A Capital Asset Pricing Model with Time Varying Covariances. *Journal of Political Economy* 95, 116–131.

Bollerslev, T., T. Law, and G. Tauchen (2008). Risk, Jumps, and Diversification. *Journal of*

Econometrics 144, 234–256.

Bollerslev, T., G. Tauchen, and H. Zhou (2009). Expected Stock Returns and Variance Risk Premia. *Review of Financial Studies* 22, 4463–4492.

Brennan, M.J., and E.S. Schwartz (1979). A Continuous-Time Approach to the Pricing Bonds. *Journal of Banking and Finance* 3, 133-155.

Britten-Jones, M. and A. Neuberger (2000). Option Prices, Implied Price Processes, and Stochastic Volatility. *Journal of Finance* 55, 839-866.

Cai, L. and N.R. Swanson (2010). An Empirical Assessment of Spot Rate Model Stability. Working Paper, Rutgers University.

Carr, P., H. Geman, D.B. Madan, and M. Yor (2002). The Fine Structure of Asset Returns: an Empirical Investigation. *Journal of Business* 75, 305–332.

Carr, P. and D. Madan (1998). Towards a Theory of Volatility Trading. Risk Books, 417-427 (Chapter 29).

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

Carr, P. and L. Wu (2009) Variance Risk Premia. *Review of Financial Studies* 22, 1311-1341.

Chacko, G. and L.M. Viceira (2003). Spectral GMM Estimation of Continuous-Time Processes. *Journal of Econometrics* 116, 259-292.

Chan, K., G. Karolyi, F. Longstaff, and A. Sanders (1992). An Empirical Comparison of Alternative Models of the Short-Term Interest Rate. *Journal of Finance* 47, 1209–27.

Chernov, M. and E. Ghysels (2002). A study Towards a Unified Approach to the Joint Estimation of Objective and Risk Neutral Measures for the Purpose of Options Valuation. *Journal of Financial Economics* 56, 407-458.

Christie, A. (1982). The Stochastic Behavior of Common Stock Variances: Value, Leverage and Interest Rate Effects. *Journal of Financial Economics* 10, 407–432.

Clark, P. K. (1973). A Subordinated Stochastic Process Model with Finite Variance for Speculative Prices. *Econometrica* 41, 135-156.

Cont, R. and C. Mancini (2007). Nonparametric Tests for Analyzing the Fine Structure of Price Fluctuations. Working Paper, Columbia University Center for Financial Engineering.

Corradi, V. (2000). Reconsidering the Continuous Time Limit of the GARCH(1,1) Process. *Journal of Econometrics* 96, 145–153.

Corradi, V., W. Distaso and N.R. Swanson (2009). Predictive Density Estimators for Daily Volatility Based on the Use of Realized Measures. *Journal of Econometrics* 150, 119–138.

Corradi, V., W. Distaso and N.R. Swanson (2010). Predictive Inference for Integrated Volatility. Working paper, Rutgers University.

Corradi, V. and N.R. Swanson (2005). Bootstrap Specification Tests for Diffusion Processes. *Journal of Econometrics* 124, 117-148.

Corradi, V. and N.R. Swanson (2010). Predictive Density Construction and Accuracy Testing with Multiple Possibly Misspecified Diffusion Models. Working Paper, Rutgers University.

Corsi, F. (2004). A Simple Long Memory Model of Realized Volatility. Working Paper, University of Southern Switzerland.

Courtadon, G. (1982). The Pricing of Options on Default-Free Bonds. *Journal of Financial and Quantitative Analysis*, 17, 75-100.

Cox, J.C., J.E. Ingersoll, and S.A. Ross (1985). A Theory of the Term Structure of Interest Rates. *Econometrica* 53, 385-407.

Cox, J.C. and S. Ross (1976). The Valuation of Options for Alternative Stochastic Processes. *Journal of Financial Economics* 3, 145-166.

Dacorogna, M.M., R. Gencay, U. Müller, R.B. Olsen, and O.B. Pictet (2001). **An Introduction to High-Frequency Finance**. Academic Press: London.

Demeterfi, K., E. Derman, M. Kamal, and J. Zou. (1999). A Guide to Volatility and Variance Swaps. *Journal of Derivatives* 6, 9–32.

Duffie, D. (2001). **Dynamic Asset Pricing Theory**. Princeton University Press: Princeton: NJ.

Duffie, D. and R. Kan. (1996). A Yield-Factor Model of Interest Rates. *Mathematical Finance* 6:379–406.

Duffie, D., J. Pan, and K. Singleton (2000). Transform Analysis and Asset Pricing for Affine Junp-Diffusions. *Econometrica* 68, 1343–1376.

Dumas, B., J. Fleming, and R. E. Whaley (1998). Implied Volatility Functions: Empirical Tests. *Journal of Finance* 53, 2059–2106.

Duong, D., and N.R. Swanson (2010). Empirical Evidence on Jumps and Large Fluctuations in Individual Stocks. Working Paper, Rutgers University.

Emanuel, D. and J. Mcbeth (1982). Further Results on the Constant Elasticity of Variance Call

Option Pricing Model. *Journal of Financial Quantitative Analysis* 17, 533-554.

Engle, R.F. (1982). Autogressive Condititional Heteroskedasticity with Esimates of Variance of U.K. Inflation. *Econometrica* 50, 987-1008.

Engle, R.F., D.M. Lilien, and R.P. Robins (1987). Estimating Time-varying Risk Premia inthe Term Structure: The ARCH-M Model. *Econometrica* 55, 391–408.

Farma, F. (1965). The Behaviour of Stock Market Prices. *Journal of Business* 38, 34-105.

Garcia, R., E. Ghysels, and E. Renault (2010). The Econometrics of Option Pricing. In Y.Ait-Sahalia, L.Hansen (eds.), Handbook of financial econometrics, North Holland, Oxford and Amsterdam.

Glosten, L.R., R. Jagannathan, and D. Runkle (1993). On the Relation Between the Expected Value and the Volatility of the Nominal Excess Return on Stocks. *Journal of Finance* 48, 1779-1801.

Granger, C.W.J. and A.P. Andersen (1978). An Introduction to Bilinear Time Series Models. Vandenhoeck and Ruprecht, Göttingen.

Hansen, P.R. and A. Lunde (2006). Realized Variance and Market Microstructure noise (with comments and rejoinder). *Journal of Business & Economic Statistics* 24, 127-218.

Harvey, A.C., E. Ruiz, and N. Shephard (1994) Multivariate Stochastic Variance Models. *Review of Economic Studies* 61, 247-264.

Heston, S.L. (1993). A closed-form Solution for Options with Stochastic Volatility with Applications to Bond and Currency Options. *Review of Financial Studies* 6, 327 343.

Heston, S.L. and S. Nandi. (2000). A Closed-form GARCH Option Valuation Model. *Review of Financial Studies* 13, 585–625.

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

Jacod, J. (2007). Statistics and High-Frequency Data. SEMSTAT Princeton University. Course Notes, Tech. Rep., Université de Paris - 6.

Jacod, J. (2008). Asymptotic Properties of Realized Power Variations and Related Functionals of Semimartingales. *Stochastic Processes and Their Applications* 118*, 517-559.*

Jiang, G.J. and R.C.A. Oomen (2008). Testing for Jumps when Asset Prices are Observed with Noise - A "Swap Variance" Approach. *Journal of Econometrics* 144, 352-370.

Jiang, G.J. and Y.S. Tian (2005). The Model-Free Implied Volatility and Its Information Content. *Review of Financial Studies* 18, 1305-1342.

Karlin, S. and H.M. Taylor (1981). A Second Course in Stochastic Processes. Academic Press, San Diego.

Kristensen, D. (2010). Nonparametric Filtering of the Realized Spot Volatility: A Kernel-based Approach. *Econometric Theory* 26, 60-93.

Lee, S. and Mykland, P.A. (2008). Jumps in Financial Markets: A New Nonparametric Test and Jump Clustering. *Review of Financial Studies* 21, 2535-2563.

Lynch, D. and N. Panigirtzoglou (2003). Option Implied and Realized Measures of Variance. Working Paper. Monetary Instruments and Markets Division, Bank of England.

Mancini, C. (2001). Disentangling the Jumps of the Diffusion in a Geometric Jumping Brownian Motion. Giornale dell'Istituto Italiano degli Attuari LXIV 19–47.

Mancini, C. (2004). Estimating the Integrated Volatility in Stochastic Volatility Models with Lévy Type Jumps. Technical report, Univ. Firenze.

Mancini, C. (2009). Non-Parametric Threshold Estimation for Models with Stochastic Diffusion Coefficient and Jumps. *Scandinavian Journal of Statistics* 36, 270-296.

Mandelbrot, B. (1963). The Variation of Certain Speculative Prices. *Journal of Business* 36, 394-419.

Marsh, T. and E. Rosenfeld (1983). Stochastic Processes for Interest Rates and Equilibrium Bond Prices. *Journal of Finance* 38, 635-646.

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

Mele, A., and F. Fornari (2006). Approximation Volatility Diffusions with CEV-ARCH Models. *Journal of Economic Dynamics and Control* 30, 931-966.

Merton, R.C. (1976). Option Pricing when Underlying Stock Returns are Discontinuous. *Journal of Financial Economics 3*, 125-144.

Müller, U.A., M.M. Dacorogna, R.D. Davé, R.B. Olsen, O.V. Puctet, and J. von Weizsäcker (1997). Volatilities of Different Time Resolutions - Analyzing the Dynamics of Market Components. *Journal of Empirical Finance* 4, 213–239.

Nelson, D.B. (1990). ARCH Models as Diffusion Approximations. *Journal of Econometrics* 45, 7-38.

Nelson, D.B. (1991). Conditional Heteroskedasticity in Asset Returns: A New Approach. *Econometrica* 59, 347-370.

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. (1992). **Stochastic Integration and Differential Equations**: A New Approach. Springer-Verlag: New York: U.S.

Schwert, G.W. (1989). Why Does Stock Market Volatility Change Over Time? *Journal of Finance* 44, 1115-1153.

Schwert, G.W. (1990). Stock Volatility and the Crash of '87. *Review of Financial Studies* 3, 77-102.

Scott, L. (1997). Pricing Stock Options in a Jump-diffusion Model with Stochastic Volatility and Interest Rates: Application of Fourier Inversion Methods. *Mathematical Finance* 7, 345–358.

Singleton, K.J. (2006). **Empirical Dynamic Asset Pricing - Model Specification and Econometric Assessment**. Princeton University Press: Princeton and Oxford.

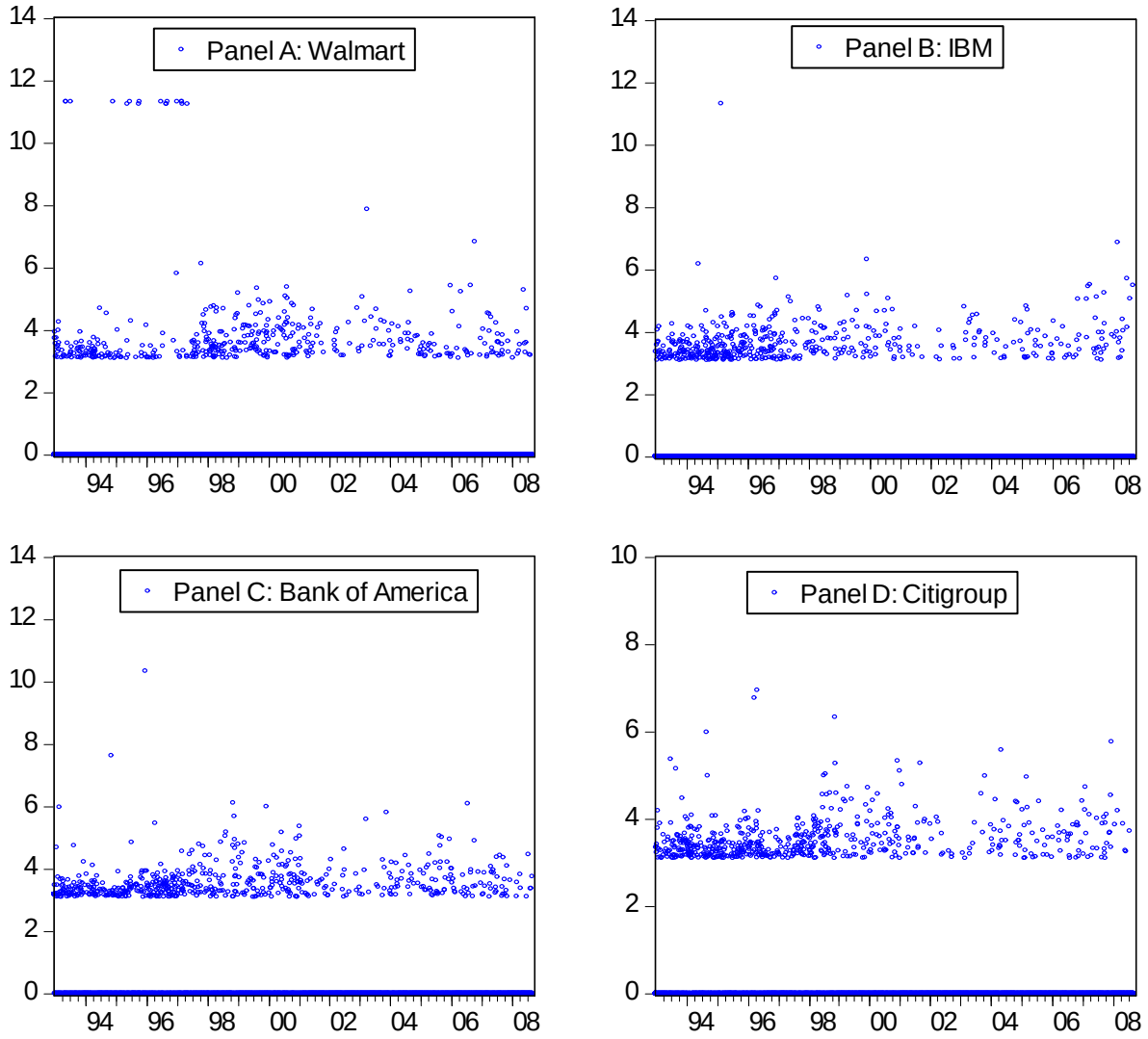
Tauchen, G. and V. Todorov (2008). Activity Signature Functions with Application for High-Frequency Data Analysis. *Journal of Econometrics* 154, 125-138.

Todorov, V. (2009). Variance Risk Premia Dynamics: The Role of Jumps. *Review of Financial Studies* 23, 345-283.

Vasicek, O. (1977). An Equilibrium Characterization of the Term Structure. *Journal of Financial Economics* 5, 177-188.

Woerner, J.H. (2006). Power and Multipower Variation: Inference for High-Frequency Data. In: **Stochastic Finance** (A. Shiryaev, M. do Rosário Grosshino, P. Oliviera and M. Esquivel, eds.) 264–276. Springer, Berlin.

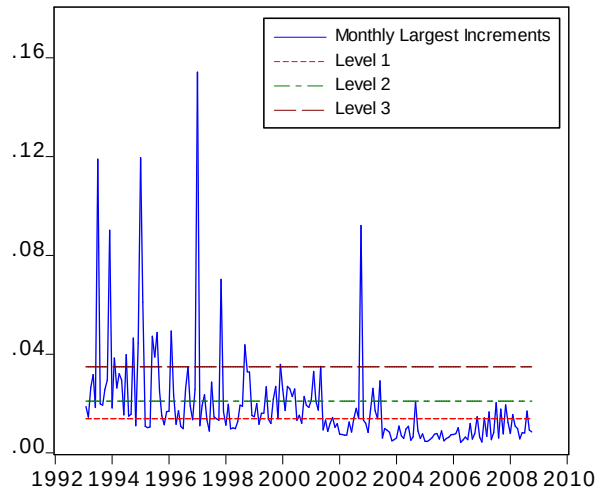
Figure 1: Jump Test Statistics of Days Identified as Having Jumps of (Log) Stock Prices: Sample Period 1993-2008 *



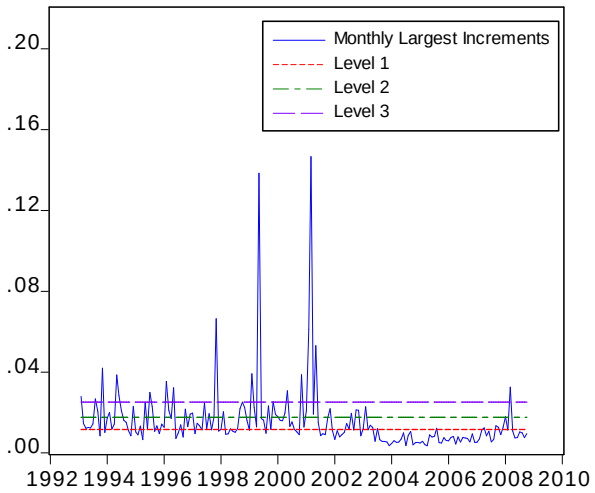
* Panel A, B, C, D depict daily test statistics of days identified as having jumps for Walmart, IBM, Bank of America, Citigroup (Log) Stock Price using 0.001 significant level. Specifically, all statistics in the figure are larger than 3.09. See section 5 for further details.

Figure 2: Monthly Largest Increments and Truncation Levels $\gamma = 1, 2, 3$: Sample Period 1993-2008 *

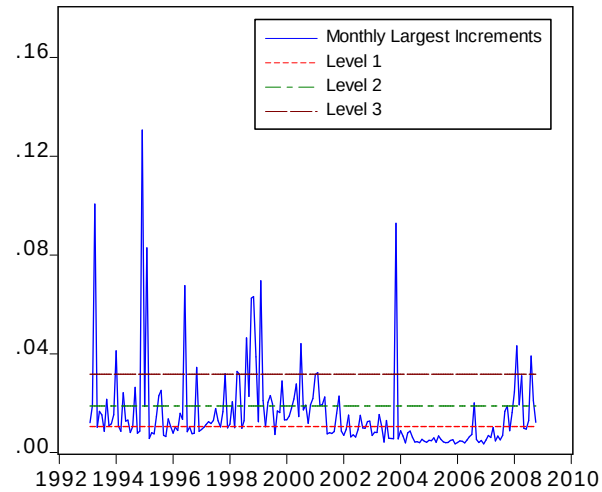
Panel A: Walmart



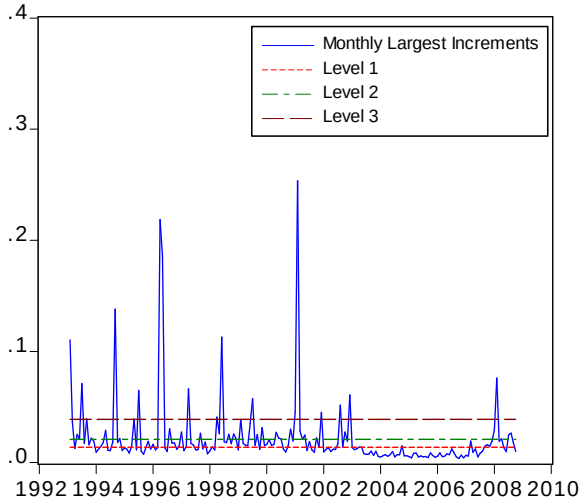
Panel B: IBM



Panel C: Bank of America



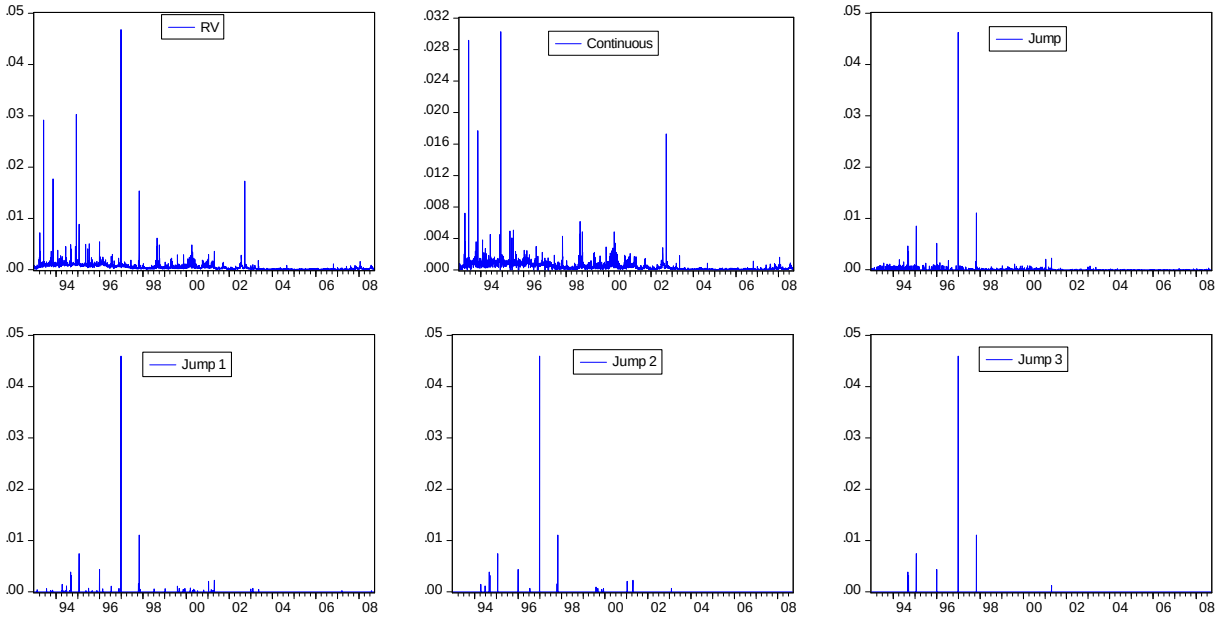
Panel D: Citigroup



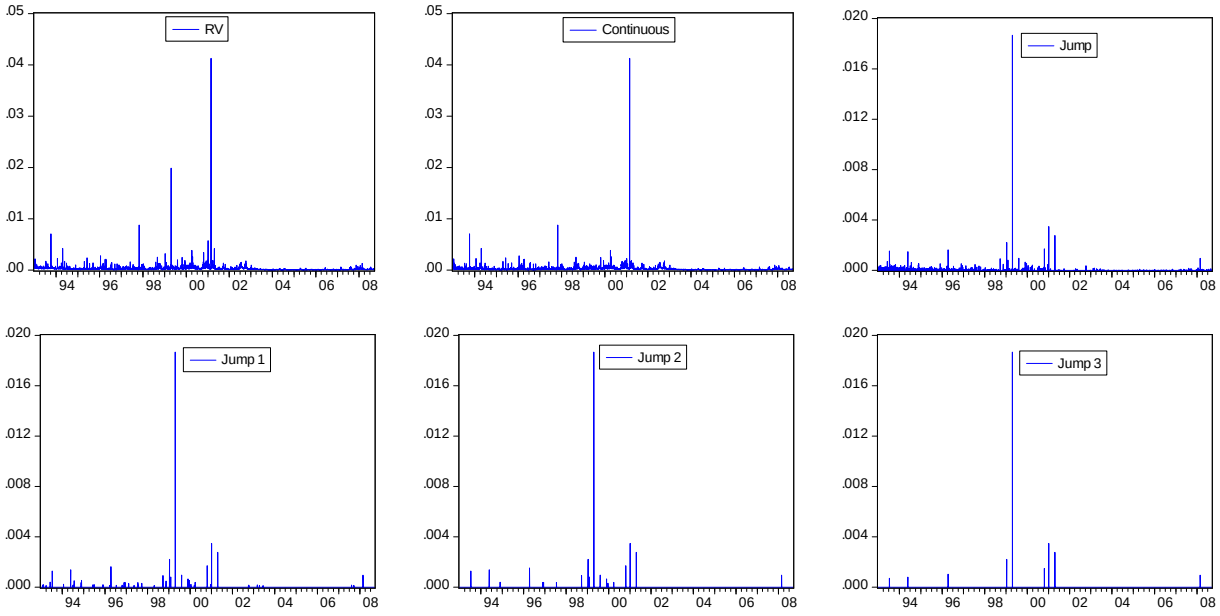
* Panel A,B,C,D depict the monthly largest absolute increments and the jump truncation levels used as thresholds in our calculations of the variations of large and small jump components, where level γ=1 corresponds to the median of monthly maximum increments , level γ=2 corresponds to 75th percentile of monthly maximum increments, and level γ=3 corresponds to 90th percentile monthly maximum increments of(log) stock prices of Walmart, IBM, Bank of America and Citigroup for the sample period is from 1993 to 2008.

Figure 3A: Daily Realized Volatility (RV) and Realized Variation of Continuous, Jump and Truncated Jump Components (Log) Stock Prices for Truncation Levels $\gamma = 1, 2, 3$ *

Panel A: Walmart



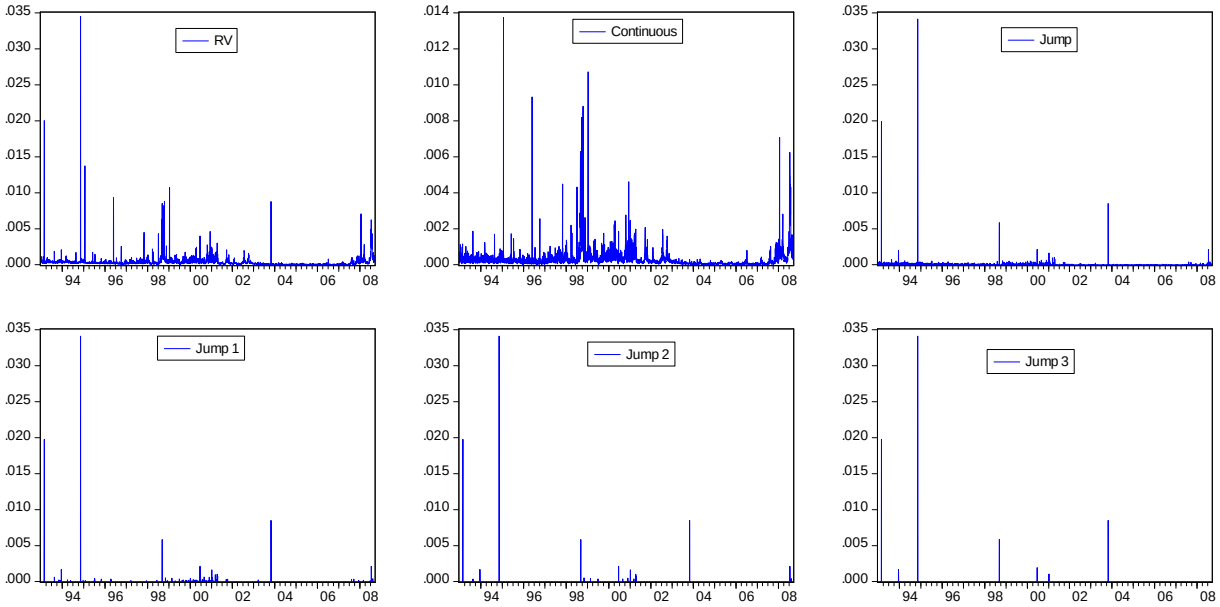
Panel B: IBM



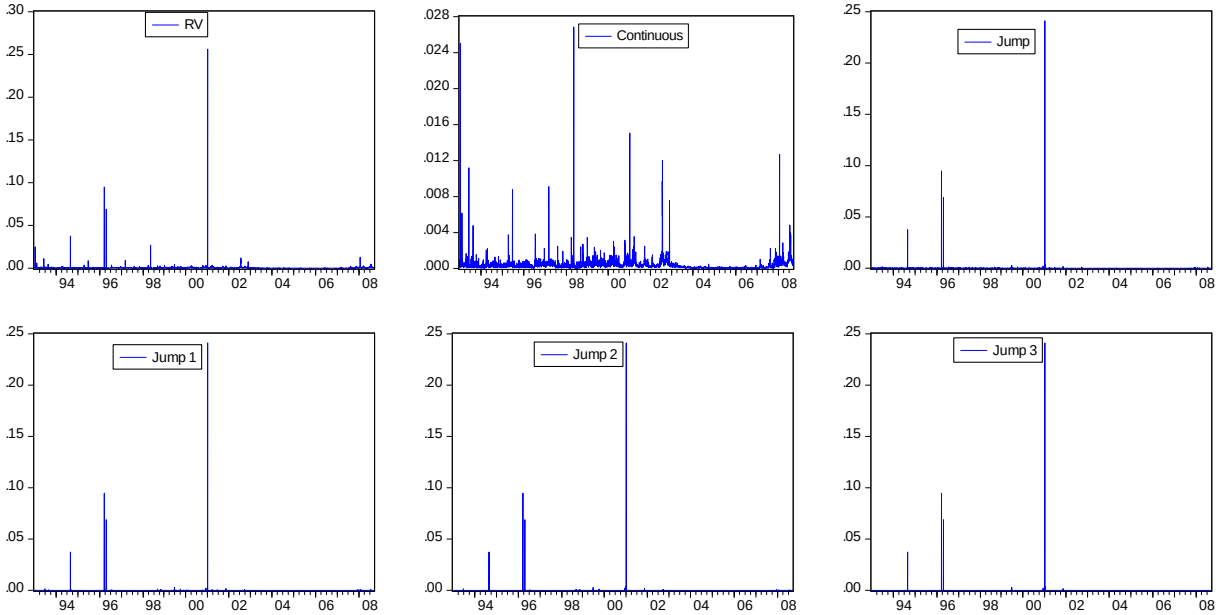
* See Figure 2 for details about the jump truncation levels. The above panels plot daily realized volatility, realized measures of the variation of continuous, jump and large jump components at truncation levels $\gamma = 1, 2, 3$, which are shortly referred to as jump 1, jump 2 and jump 3 for the period 1993-2008. The realized measures of variations are calculated as discussed in Section 4 and 5.

Figure 3B: Daily Realized Volatility (RV) and Realized Variation of Continuous, Jump and Truncated Jump Components of (Log) Stock Prices for Truncation Levels $\gamma = 1, 2, 3$ *

Panel C: Bank of America



Panel D: Citigroup



* See notes in Figure 3A.

Table 1: Percentage of Days Identified as Having Jumps Using Daily Statistics *

Stock Name	*Panel A: Sample Period 1993-2000 ($T \simeq 2000$)*						*Panel B: Sample Period 2001-2008 ($T \simeq 1900$)*					
Significant Level α	0.1	0.05	0.01	0.005	0.001	0.0001	0.1	0.05	0.01	0.005	0.001	0.0001
Alcoa	88	80.9	62.1	52	26.3	8.7	49.2	39.3	21.4	16.5	9.7	4.3
American Express	82.9	76.2	56.4	46.6	19.3	4.6	47.4	36.8	20.2	14.7	8.0	4.0
Bank of America	81.8	75.6	56.7	46.4	21.3	5.0	45.1	34.1	20.1	15.7	8.8	3.1
Citigroup	86.3	80.5	63.3	51.9	23.3	4.9	43.6	32.9	17.9	14.6	7.1	2.6
Caterpillar	87.2	81.5	61.8	51	25.9	7.3	46	35.3	19.9	16.3	9.5	4.3
Dupont	83.8	76.5	57.2	48.3	24.2	5.6	49.5	38.8	21.8	17.1	9.5	3.9
Walt Disney	89.3	83.9	65.9	56.0	27.3	5.3	55.6	43.9	23.9	17.6	10.1	3.9
General Electric	79.6	73.5	54.5	45.5	22.3	4.5	49.2	39.3	21.8	16.2	9.4	3.9
GM	88.1	83.1	65.4	54	25.4	6.2	51.8	40.4	22.8	17.8	10.5	4.7
Home Depot	87.7	81	62.1	51.4	24.6	5.1	49.5	38.5	22.1	16.8	10	4.3
IBM	73.8	65	47.3	39.6	19.2	5.9	39.9	30.1	15.1	11.7	5.9	2.8
Intel	69.2	58.9	39.5	33.0	18.0	6.3	51.7	41.4	23.6	18.7	11.3	4.7
Johnson & Johnson	86.7	81.2	62.8	52.5	25.2	5.7	47.5	37.7	22.1	18.0	10.9	4.6
JPM	79.5	73.2	55.7	47.6	21.4	5.0	47.9	35.9	20.8	16.1	9.0	3.3
Coca Cola	86.4	80.8	63.3	54.2	23.9	4.8	52.5	41.9	23.3	18.5	10.2	4.6
McDonald's	90.5	85	66.1	55.9	25.8	4.9	51.3	40.8	24.6	19.8	11.5	4.8
3M	85.7	78.8	59.2	49.9	25.6	6.9	43.1	33.1	18.8	14.2	7.9	3.6
Microsoft	68.5	58.7	38.6	30.5	16.4	7.0	56.3	44.8	25.7	21.5	11.1	4.4
Pfizer	82.6	75.4	56.6	49.1	26.3	6.5	50	40	23.5	17.7	9.4	4.1
Procter & Gamble	80.1	72.4	55.6	46.4	25.5	6.4	46.9	35.6	18.5	14.4	7.2	2.8
AT & T	89.3	83.3	65.8	54.7	23.1	4.4	58.8	48.4	29	22.8	13.8	6.1
United Tech.Corp.	84.2	77.1	54.3	43.9	22.8	8.2	46.3	36.3	20.5	16.0	9.1	3.6
Verizon	81.5	67.7	46	39.5	24.2	8.1	51.4	40.9	24.5	19.4	11.2	5.0
Walmart	86.7	81.5	59.8	46.9	15.5	5.1	44.7	34.3	18.7	14.0	7.4	2.6
ExxonMobil	61.3	49.8	32.8	26.2	17	5.2	44.2	33.6	17.5	12.9	6.2	2.9
Average	82.4	75.3	56.4	46.9	22.8	5.9	48.8	38.2	21.5	16.8	9.4	4.0

* See notes to Figure 1. Entries denote the percentage of days identified as having jumps based on the calculation of daily statistics. Statistics are the adjusted ratio jump statistics of Barndorff-Nielsen and Shephard (2006) and Huang and Tauchen (2005), as discussed in Section 4. Test results are summarized in Panel A for the sample period from 1993-2000 and for the sample period 2001-2008 in Panel B. These sample periods have approximately 2000 and 1900 daily statistics, respectively. Statistics are reported for six different significance levels, $\alpha = 0.1, 0.05, 0.01, 0.005, 0.001, 0.0001$.

Table 2: Daily Realized Variation: Ratio of Jump to Total Variation *

Stock Name	*Panel A: Sample Period 1993-2000 ($T \simeq 2000$)*						*Panel B: Sample Period 2001-2008 ($T \simeq 1900$)*					
Significant Level α	0.1	0.05	0.01	0.005	0.001	0.0001	0.1	0.05	0.01	0.005	0.001	0.0001
Alcoa	41.2	39.5	32.6	28.4	14.8	5.2	13.6	11.8	7.7	6.4	4.1	2.1
American Express	39.5	38.1	30.5	26.2	10.8	2.6	12.7	10.9	7.1	5.7	3.5	1.9
Bank of America	37.2	36.0	29.7	25.3	12.1	2.9	11.6	9.8	6.8	5.6	3.5	1.5
Citigroup	41.5	40.3	34.1	29.2	13.7	3.0	11.0	9.3	6.0	5.1	2.8	1.2
Caterpillar	40.0	38.6	32.0	27.5	14.7	4.5	12.3	10.5	7.1	6.2	4.0	2.0
Dupont	36.8	35.4	29.1	25.8	13.3	3.2	13.3	11.5	7.6	6.4	4.0	1.8
Walt Disney	42.0	40.8	34.9	30.7	15.8	3.4	15.3	13.2	8.6	6.9	4.2	1.8
General Electric	34.0	32.8	27.1	23.8	12.3	2.5	13.2	11.6	7.7	6.1	3.8	1.8
GM	42.5	41.4	35.2	30.2	14.7	3.9	14.6	12.5	8.3	6.9	4.4	2.2
Home Depot	40.0	38.6	32.3	27.9	14.1	3.1	13.3	11.5	7.8	6.3	4.1	2.0
IBM	30.1	28.3	23.1	20.4	10.7	3.5	9.8	8.2	5.1	4.2	2.3	1.3
Intel	24.0	22.0	16.8	14.7	8.9	3.5	13.9	12.2	8.3	6.9	4.6	2.2
Johnson &Johnson	39.2	38.1	32.3	28.3	14.2	3.4	13.0	11.3	7.8	6.7	4.4	2.1
JPM	36.0	34.7	29.2	25.6	12.0	2.8	12.6	10.6	7.2	5.9	3.7	1.6
Coca Cola	39.9	38.8	33.1	29.4	13.7	2.8	14.0	12.3	8.2	6.8	4.2	2.1
McDonald's	43.9	42.8	36.0	31.5	15.1	3.1	14.9	13.1	9.2	7.8	5.0	2.3
3M	39.0	37.5	30.9	27.2	14.4	4.1	11.0	9.4	6.2	5.1	3.2	1.6
Microsoft	23.2	21.4	16.1	13.5	8.0	3.7	15.1	13.2	9.0	7.8	4.5	2.0
Pfizer	35.5	34.1	28.2	25.4	14.7	3.9	13.8	12.1	8.3	6.7	3.9	1.9
Procter &Gamble	33.9	32.4	27.8	24.5	14.2	3.7	11.9	10.1	6.3	5.2	3.0	1.4
AT &T	43.8	42.6	36.0	31.1	13.5	2.8	17.7	15.9	11.1	9.3	6.1	3.0
United Tech.Corp.	37.1	35.5	27.6	23.0	12.4	4.8	12.2	10.5	7.0	5.8	3.6	1.6
Verizon	29.4	26.8	21.0	18.8	12.0	4.4	14.4	12.6	8.7	7.4	4.7	2.4
Walmart	44.4	43.2	33.8	27.4	9.1	3.4	11.5	9.8	6.3	5.1	3.0	1.2
ExxonMobil	19.3	17.1	12.7	10.9	7.6	2.6	10.8	9.1	5.7	4.6	2.5	1.3
Average	36.5	35.1	28.9	25.1	12.7	3.5	13.1	11.3	7.6	6.3	3.9	1.9

Panel C: Sample Period 1993-2008 ($T \simeq 3900$)

Significant Level α	0.1	0.05	0.01	0.005	0.001	0.0001
Alcoa	26.7	25.0	19.5	16.9	9.2	3.6
American Express	26.4	24.8	19.1	16.2	7.2	2.3
Bank of America	24.7	23.2	18.5	15.7	7.9	2.2
Citigroup	26.6	25.1	20.3	17.4	8.3	2.1
Caterpillar	26.2	24.7	19.6	16.9	9.4	3.3
Dupont	25.4	23.7	18.6	16.3	8.8	2.5
Walt Disney	29.0	27.3	22.0	19.1	10.2	2.6
General Electric	23.8	22.5	17.6	15.1	8.2	2.2
GM	28.9	27.3	22.1	18.8	9.7	3.1
Home Depot	27.0	25.3	20.3	17.4	9.2	2.5
IBM	20.2	18.5	14.3	12.5	6.6	2.4
Intel	19.0	17.2	12.6	10.9	6.8	2.9
Johnson &Johnson	26.4	25.0	20.3	17.7	9.4	2.8
JPM	24.6	23.0	18.4	16.0	7.9	2.2
Coca Cola	27.3	25.9	20.9	18.4	9.0	2.5
McDonald's	30.4	29.0	23.5	20.5	10.4	2.7
3M	26.0	24.4	19.4	16.9	9.2	2.9
Microsoft	19.4	17.6	12.8	10.8	6.3	2.9
Pfizer	24.9	23.3	18.5	16.3	9.4	2.9
Procter &Gamble	23.2	21.5	17.3	15.1	8.7	2.6
AT &T	31.2	29.6	23.9	20.5	9.9	2.9
United Tech.Corp.	23.7	22.1	16.5	13.8	7.7	3.1
Verizon	15.3	13.5	9.5	8.1	5.1	2.5
Walmart	28.3	26.9	20.4	16.5	6.1	2.3
ExxonMobil	11.9	10.1	6.6	5.4	3.1	1.5
Average	24.7	23.1	18.1	15.6	8.1	2.6

* See notes to Figure 2. The entries in the table denote the average percentage of daily variation of the jump component relative to daily realized variance for the sample periods 1993-2000, 2001-2008 and 1993-2008. The realized measure of variation of the jump component is calculated as discussed in Section 4. In addition to frequency of jumps, realized measures of variations also take the magintude of jumps into account. Entries are caculcated accross 6 different significant levels, $\alpha = 0.1, 0.05, 0.01, 0.005, 0.001, 0.0001$.

Table 3A: Daily Realized Variation: Ratio of Large Jump to Total Variation, Jump Truncation Level $\gamma = 1$ *

Stock Name	*Panel A: Sample Period 1993-2000 ($T \simeq 2000$)*						*Panel B: Sample Period 2001-2008 ($T \simeq 1900$)*					
Significant Level α	0.1	0.05	0.01	0.005	0.001	0.0001	0.1	0.05	0.01	0.005	0.001	0.0001
Alcoa	1.8	1.7	1.2	1.1	0.9	0.6	1.3	1.2	0.9	0.8	0.5	0.4
American Express	1.5	1.3	1.0	0.8	0.6	0.4	1.2	1.1	0.8	0.6	0.5	0.3
Bank of America	3.0	2.7	1.8	1.5	1.0	0.7	1.0	0.9	0.8	0.6	0.4	0.2
Citigroup	2.1	1.9	1.2	1.0	0.6	0.4	0.8	0.7	0.5	0.4	0.3	0.1
Caterpillar	2.3	2.2	1.6	1.5	1.0	0.6	0.7	0.6	0.5	0.5	0.3	0.2
Dupont	2.1	1.9	1.2	1.0	0.7	0.3	0.9	0.9	0.6	0.5	0.4	0.2
Walt Disney	2.1	1.8	1.1	0.9	0.6	0.4	1.6	1.4	1.0	0.9	0.6	0.3
General Electric	1.5	1.4	0.8	0.7	0.4	0.2	1.3	1.2	0.9	0.6	0.3	0.2
GM	1.5	1.4	1.0	0.8	0.6	0.4	1.3	1.2	0.8	0.7	0.5	0.2
Home Depot	1.9	1.7	1.3	1.1	0.6	0.3	0.7	0.6	0.5	0.3	0.2	0.1
IBM	2.3	2.1	1.7	1.6	1.0	0.7	0.5	0.5	0.4	0.4	0.2	0.1
Intel	2.3	2.1	1.6	1.2	0.8	0.5	0.7	0.7	0.4	0.4	0.3	0.2
Johnson &Johnson	2.2	2.0	1.5	1.3	0.9	0.6	0.6	0.6	0.4	0.3	0.2	0.1
JPM	1.3	1.1	0.7	0.6	0.3	0.2	1.7	1.5	1.0	0.8	0.6	0.3
Coca Cola	2.3	2.1	1.3	1.2	0.8	0.5	0.8	0.8	0.6	0.5	0.4	0.3
McDonald's	1.8	1.6	1.2	0.9	0.7	0.4	1.0	1.0	0.7	0.6	0.4	0.2
3M	2.1	2.0	1.3	1.1	0.7	0.5	0.6	0.6	0.4	0.4	0.4	0.3
Microsoft	2.9	2.7	2.0	1.7	1.0	0.5	0.6	0.6	0.5	0.4	0.2	0.1
Pfizer	2.0	1.9	1.3	1.1	0.8	0.5	0.7	0.6	0.5	0.4	0.3	0.2
Procter &Gamble	2.4	2.2	1.7	1.4	0.9	0.5	0.8	0.7	0.5	0.4	0.3	0.2
AT &T	2.3	2.2	1.6	1.3	0.9	0.7	1.7	1.5	1.2	1.1	0.7	0.4
United Tech.Corp.	3.2	2.9	2.1	1.9	1.2	0.6	1.3	1.1	0.8	0.7	0.5	0.2
Verizon	6.9	6.3	5.0	4.4	2.7	1.1	1.5	1.3	1.0	0.9	0.6	0.4
Walmart	2.7	2.4	1.4	1.2	0.9	0.6	0.6	0.6	0.4	0.4	0.3	0.1
ExxonMobil	2.1	1.8	1.3	1.0	0.6	0.4	1.1	0.9	0.7	0.6	0.4	0.2
Average	2.3	2.1	1.5	1.3	0.8	0.5	1.0	0.9	0.7	0.6	0.4	0.2

Panel C: Sample Period 1993-2008 ($T \simeq 3900$)

Significant Level α	0.1	0.05	0.01	0.005	0.001	0.0001
Alcoa	1.6	1.4	1.0	0.9	0.7	0.5
American Express	1.3	1.2	0.9	0.7	0.5	0.4
Bank of America	2.0	1.9	1.3	1.1	0.7	0.5
Citigroup	1.4	1.3	0.8	0.7	0.5	0.3
Caterpillar	1.5	1.4	1.1	1.0	0.7	0.4
Dupont	1.5	1.4	0.9	0.8	0.5	0.2
Walt Disney	1.8	1.6	1.1	0.9	0.6	0.4
General Electric	1.4	1.3	0.8	0.7	0.4	0.2
GM	1.4	1.3	0.9	0.8	0.5	0.3
Home Depot	1.3	1.2	0.9	0.7	0.4	0.2
IBM	1.4	1.3	1.0	1.0	0.6	0.4
Intel	1.5	1.4	1.0	0.8	0.6	0.3
Johnson &Johnson	1.4	1.3	1.0	0.8	0.6	0.4
JPM	1.5	1.3	0.9	0.7	0.5	0.3
Coca Cola	1.6	1.5	1.0	0.8	0.6	0.4
McDonald's	1.4	1.3	1.0	0.8	0.5	0.3
3M	1.4	1.3	0.9	0.8	0.6	0.4
Microsoft	1.8	1.7	1.3	1.1	0.6	0.3
Pfizer	1.3	1.2	0.9	0.8	0.6	0.4
Procter &Gamble	1.6	1.5	1.1	0.9	0.6	0.3
AT &T	2.0	1.9	1.4	1.2	0.8	0.5
United Tech.Corp.	2.1	2.0	1.4	1.3	0.8	0.4
Verizon	1.8	1.6	1.2	1.1	0.7	0.4
Walmart	1.7	1.5	0.9	0.8	0.6	0.3
ExxonMobil	1.2	1.1	0.7	0.6	0.4	0.2
Average	1.6	1.4	1.0	0.9	0.6	0.3

* See notes to Figure 2. Entries in the table denote the average percentage of daily variation due to jumps constructed using truncation level $\gamma = 1$, relative to the daily realized variance, for the sample periods 1993-2000, 2001-2008 and 1993-2008. The realized measure of variation of the jump component is calculated as discussed in Section 4. Entries are caculcated accross 6 different significance levels (α =0.1, 0.05, 0.01, 0.005, 0.001, 0.0001).

Table 3B: Daily Realized Variation: Ratio of Large Jump to Total Variation, Jump Truncation Level $\gamma = 2$ *

Stock Name	*Panel A: Sample Period 1993-2000 ($T \simeq 2000$)*						*Panel B: Sample Period 2001-2008 ($T \simeq 1900$)*					
Significant Level α	0.1	0.05	0.01	0.005	0.001	0.0001	0.1	0.05	0.01	0.005	0.001	0.0001
Alcoa	0.8	0.8	0.6	0.5	0.5	0.3	0.5	0.5	0.4	0.4	0.2	0.1
American Express	0.8	0.7	0.6	0.5	0.3	0.3	0.5	0.4	0.3	0.2	0.2	0.1
Bank of America	0.9	0.9	0.5	0.4	0.3	0.2	0.4	0.4	0.3	0.3	0.2	0.1
Citigroup	1.0	1.0	0.7	0.5	0.4	0.3	0.3	0.3	0.2	0.2	0.2	0.1
Caterpillar	1.0	1.0	0.8	0.8	0.5	0.3	0.3	0.3	0.3	0.3	0.2	0.2
Dupont	0.9	0.8	0.4	0.3	0.2	0.1	0.4	0.4	0.2	0.2	0.1	0.0
Walt Disney	1.0	0.9	0.4	0.3	0.3	0.2	0.5	0.5	0.3	0.3	0.2	0.1
General Electric	0.7	0.7	0.4	0.4	0.2	0.1	0.6	0.6	0.5	0.3	0.1	0.1
GM	0.7	0.7	0.4	0.3	0.2	0.2	0.7	0.6	0.4	0.4	0.3	0.1
Home Depot	1.0	0.9	0.7	0.6	0.3	0.2	0.2	0.2	0.2	0.1	0.0	0.0
IBM	1.0	0.9	0.8	0.7	0.5	0.3	0.2	0.2	0.2	0.2	0.1	0.1
Intel	0.9	0.9	0.7	0.6	0.4	0.2	0.2	0.2	0.1	0.1	0.1	0.0
Johnson &Johnson	0.9	0.8	0.6	0.5	0.4	0.3	0.3	0.3	0.2	0.2	0.2	0.1
JPM	0.4	0.4	0.2	0.2	0.1	0.0	0.7	0.7	0.5	0.4	0.3	0.2
Coca Cola	0.9	0.9	0.4	0.4	0.2	0.2	0.4	0.4	0.2	0.2	0.2	0.2
McDonald's	0.8	0.7	0.5	0.4	0.3	0.3	0.5	0.5	0.4	0.2	0.2	0.0
3M	1.0	0.9	0.5	0.4	0.3	0.2	0.2	0.1	0.1	0.1	0.1	0.1
Microsoft	1.1	1.0	0.8	0.6	0.3	0.1	0.2	0.2	0.2	0.2	0.1	0.0
Pfizer	0.8	0.8	0.6	0.6	0.4	0.3	0.3	0.3	0.2	0.2	0.1	0.1
Procter &Gamble	1.1	1.0	0.8	0.6	0.4	0.3	0.3	0.2	0.2	0.2	0.1	0.1
AT &T	1.1	1.0	0.7	0.6	0.4	0.4	0.7	0.6	0.6	0.5	0.3	0.1
United Tech.Corp.	1.1	1.0	0.7	0.6	0.5	0.3	0.4	0.4	0.3	0.3	0.2	0.1
Verizon	2.8	2.6	2.2	1.9	0.7	0.2	0.5	0.5	0.4	0.4	0.3	0.2
Walmart	1.2	1.1	0.6	0.5	0.5	0.4	0.2	0.2	0.2	0.2	0.1	0.0
ExxonMobil	0.8	0.6	0.5	0.4	0.4	0.2	0.5	0.5	0.4	0.3	0.2	0.1
Average	1.0	0.9	0.6	0.5	0.4	0.2	0.4	0.4	0.3	0.3	0.2	0.1

Panel C: Sample Period 1993-2008 ($T \simeq 3900$)

Significant Level α	0.1	0.05	0.01	0.005	0.001	0.0001
Alcoa	0.6	0.6	0.5	0.4	0.3	0.2
American Express	0.7	0.6	0.4	0.3	0.3	0.2
Bank of America	0.7	0.6	0.4	0.4	0.3	0.2
Citigroup	0.7	0.7	0.5	0.4	0.3	0.2
Caterpillar	0.7	0.6	0.5	0.5	0.4	0.3
Dupont	0.7	0.6	0.3	0.3	0.2	0.1
Walt Disney	0.8	0.7	0.4	0.3	0.2	0.2
General Electric	0.7	0.6	0.5	0.3	0.1	0.1
GM	0.7	0.7	0.4	0.3	0.2	0.2
Home Depot	0.6	0.6	0.4	0.3	0.2	0.1
IBM	0.6	0.6	0.5	0.5	0.3	0.2
Intel	0.6	0.5	0.4	0.3	0.2	0.1
Johnson &Johnson	0.6	0.6	0.4	0.4	0.3	0.2
JPM	0.6	0.5	0.4	0.3	0.2	0.1
Coca Cola	0.6	0.6	0.3	0.3	0.2	0.2
McDonald's	0.7	0.6	0.5	0.3	0.3	0.2
3M	0.6	0.6	0.3	0.3	0.2	0.2
Microsoft	0.6	0.6	0.5	0.4	0.2	0.0
Pfizer	0.6	0.5	0.4	0.4	0.3	0.2
Procter &Gamble	0.7	0.6	0.5	0.4	0.3	0.2
AT &T	0.9	0.8	0.6	0.6	0.4	0.3
United Tech.Corp.	0.7	0.7	0.5	0.4	0.3	0.2
Verizon	0.7	0.6	0.5	0.4	0.3	0.2
Walmart	0.7	0.6	0.4	0.3	0.3	0.2
ExxonMobil	0.6	0.5	0.4	0.3	0.2	0.1
Average	0.7	0.6	0.4	0.4	0.3	0.2

* See notes to Table 3A.

Table 3C: Daily Realized Variation: Ratio of Large Jump to Total Variation, Jump Truncation Level $\gamma = 3$ *

Stock Name	*Panel A: Sample Period 1993-2000 ($T \simeq 2000$)*						*Panel B: Sample Period 2001-2008 ($T \simeq 1900$)*					
Significant Level α	0.1	0.05	0.01	0.005	0.001	0.0001	0.1	0.05	0.01	0.005	0.001	0.0001
Alcoa	0.3	0.3	0.2	0.2	0.1	0.1	0.2	0.2	0.2	0.1	0.1	0.0
American Express	0.3	0.3	0.2	0.2	0.2	0.2	0.2	0.2	0.2	0.1	0.1	0.0
Bank of America	0.5	0.5	0.3	0.2	0.2	0.2	0.1	0.1	0.1	0.1	0.1	0.1
Citigroup	0.5	0.5	0.4	0.3	0.2	0.2	0.2	0.2	0.1	0.1	0.1	0.0
Caterpillar	0.5	0.5	0.4	0.4	0.3	0.2	0.1	0.1	0.1	0.1	0.1	0.1
Dupont	0.4	0.4	0.2	0.1	0.1	0.0	0.1	0.1	0.0	0.0	0.0	0.0
Walt Disney	0.5	0.5	0.2	0.2	0.2	0.2	0.3	0.3	0.2	0.2	0.1	0.0
General Electric	0.3	0.3	0.2	0.1	0.0	0.0	0.3	0.3	0.3	0.2	0.0	0.0
GM	0.3	0.3	0.2	0.1	0.1	0.1	0.2	0.2	0.1	0.1	0.1	0.0
Home Depot	0.4	0.4	0.3	0.3	0.1	0.1	0.0	0.0	0.0	0.0	0.0	0.0
IBM	0.3	0.3	0.3	0.2	0.2	0.1	0.1	0.1	0.1	0.1	0.1	0.1
Intel	0.4	0.4	0.3	0.2	0.1	0.1	0.1	0.1	0.0	0.0	0.0	0.0
Johnson &Johnson	0.4	0.4	0.2	0.2	0.2	0.2	0.2	0.2	0.2	0.1	0.1	0.0
JPM	0.2	0.2	0.1	0.1	0.0	0.0	0.2	0.2	0.2	0.2	0.1	0.1
Coca Cola	0.5	0.5	0.2	0.2	0.1	0.1	0.2	0.2	0.1	0.1	0.1	0.1
McDonald's	0.4	0.3	0.3	0.2	0.2	0.2	0.2	0.2	0.1	0.1	0.1	0.0
3M	0.3	0.3	0.2	0.1	0.1	0.1	0.1	0.1	0.1	0.0	0.0	0.0
Microsoft	0.4	0.4	0.2	0.2	0.1	0.0	0.1	0.1	0.1	0.1	0.0	0.0
Pfizer	0.3	0.3	0.2	0.2	0.2	0.2	0.1	0.1	0.1	0.1	0.0	0.0
Procter &Gamble	0.3	0.3	0.2	0.2	0.2	0.1	0.2	0.2	0.1	0.1	0.1	0.1
AT &T	0.6	0.5	0.3	0.3	0.2	0.2	0.2	0.2	0.2	0.1	0.1	0.0
United Tech.Corp.	0.2	0.2	0.2	0.2	0.2	0.1	0.2	0.2	0.1	0.1	0.1	0.0
Verizon	0.7	0.7	0.7	0.7	0.0	0.0	0.2	0.1	0.1	0.1	0.1	0.0
Walmart	0.5	0.5	0.3	0.2	0.2	0.2	0.1	0.1	0.1	0.1	0.0	0.0
ExxonMobil	0.1	0.1	0.1	0.0	0.0	0.0	0.2	0.2	0.2	0.2	0.1	0.0
Average	0.4	0.4	0.3	0.2	0.1	0.1	0.2	0.2	0.1	0.1	0.1	0.0

Panel C: Sample Period 1993-2008 ($T \simeq 3900$)						
Significant Level α	0.1	0.05	0.01	0.005	0.001	0.0001
Alcoa	0.3	0.3	0.2	0.2	0.1	0.1
American Express	0.3	0.2	0.2	0.1	0.1	0.1
Bank of America	0.3	0.3	0.2	0.2	0.1	0.1
Citigroup	0.3	0.3	0.2	0.2	0.2	0.1
Caterpillar	0.3	0.3	0.3	0.2	0.2	0.1
Dupont	0.3	0.3	0.1	0.0	0.0	0.0
Walt Disney	0.4	0.4	0.2	0.2	0.1	0.1
General Electric	0.3	0.3	0.2	0.2	0.0	0.0
GM	0.3	0.3	0.1	0.1	0.1	0.1
Home Depot	0.2	0.2	0.2	0.1	0.0	0.0
IBM	0.2	0.2	0.2	0.2	0.1	0.1
Intel	0.2	0.2	0.2	0.1	0.1	0.0
Johnson &Johnson	0.3	0.3	0.2	0.2	0.2	0.1
JPM	0.2	0.2	0.1	0.1	0.0	0.0
Coca Cola	0.3	0.3	0.1	0.1	0.1	0.1
McDonald's	0.3	0.3	0.2	0.1	0.1	0.1
3M	0.2	0.2	0.1	0.1	0.1	0.1
Microsoft	0.2	0.2	0.2	0.1	0.1	0.0
Pfizer	0.2	0.2	0.2	0.1	0.1	0.1
Procter &Gamble	0.2	0.2	0.2	0.2	0.1	0.1
AT &T	0.4	0.4	0.2	0.2	0.2	0.1
United Tech.Corp.	0.2	0.2	0.2	0.2	0.1	0.1
Verizon	0.2	0.2	0.1	0.1	0.1	0.0
Walmart	0.3	0.3	0.2	0.2	0.1	0.1
ExxonMobil	0.2	0.2	0.2	0.2	0.1	0.0
Average	0.3	0.3	0.2	0.1	0.1	0.1

* See notes to Table 3A.