

# Time-varying Volatility and the Power Law Distribution of Stock Returns

Missaka Warusawitharana\*

Board of Governors of the Federal Reserve System

August 5, 2018

JEL Classifications: C58, D30, G12

Keywords: Tail distributions, high frequency returns, power laws, time-varying volatility

---

\*I thank Yacine Aït-Sahalia, Dobislav Dobrev, Joshua Gallin, Tugkan Tuzun, Toni Whited, Jonathan Wright, Yesol Yeh, an anonymous referee and seminar participants at the Federal Reserve Board and the 2017 Computing in Economics and Finance conference for helpful comments on the paper. Contact: Division of Research and Statistics, Board of Governors of the Federal Reserve System, Mail Stop 97, 20th and C Street NW, Washington, D.C. 20551. [missaka.n.warusawitharana@frb.gov](mailto:missaka.n.warusawitharana@frb.gov), (202)452-3461.

## **Abstract**

While many studies find that the tail distribution of high frequency stock returns follows a power law, there are only a few explanations for this finding. This study presents evidence that time-varying volatility can account for the power law property of high frequency stock returns. In particular, one finds that a conditional normal model with nonparametric volatility provides a strong fit. Specifically, a cross-sectional regression of the power law coefficients obtained from stock returns on the coefficients implied by the nonparametric volatility model yields a slope close to one. Further, for most of the stocks in the sample taken individually, the model-implied coefficient falls within the 95 percent confidence interval for the coefficient estimated from returns data.

# 1 Introduction

A growing literature has documented that the tail distributions of a broad range of data in the natural and social sciences follow a power law, which implies that the tail probability of a series declines as a power function as it increases in value.<sup>1</sup> This finding has drawn attention both due to the simplicity and scalability of the power law as well as for the ubiquity of this relationship across many fields and data sets. As noted by Stumpf and Porter (2012), however, our understanding of the causes and mechanisms that underpin many of these relationships lag the empirical evidence on them. This study examines whether time-varying volatility can help explain one such finding—the power law property of the tail distribution of high-frequency stock returns documented by Plerou, Gopikrishnan, Amaral, Meyer, and Stanley (1999), among others.

Gabaix, Gopikrishnan, Plerou, and Stanley (2003, 2006) provide the leading explanation for the power law property of stock returns. These studies argue that the tail distribution of stock returns arises from the price impact of trades initiated by different market participants, who themselves have a tail distribution of assets. Using data on the tail distributions of trades and assets, they demonstrate how a model of price impact can simultaneously explain the tail distributions of trades and returns, as these processes in the model inherit the power law property of investor assets.

Building on the insight of Clark (1973) and others that mixture distributions have heavy tails, this study examines whether time-varying volatility can generate the power law property. The key insight underlying the study is that a conditional normal model with time-varying volatility may exhibit a power law distribution in the tails. This result is first established using a stylized conditional normal model with an inverse-gamma distribution for return variance. Furthermore, the study shows that any thin-tailed distribution for conditional returns combined with a power law for the tail distribution of volatility yields the power law property for returns. The empirical analysis begins with the stylized conditional

---

<sup>1</sup>Gopikrishnan, Plerou, Gabaix, and Stanley (2000) find that tail distribution of returns on stock indices exhibit a power law with a power coefficient around 3; Plerou, Gopikrishnan, Amaral, Meyer, and Stanley (1999) find that the tail distribution returns on individual stocks follow a power law with coefficients ranging from 2.5 to 4; Axtell (2001) shows that the firm size distribution follows a power law with a coefficient of 1; Rozenfeld, Rybski, Gabaix, and Makse (2011) present evidence that the distribution of populated areas follows a power law with a coefficient of 1; Toda (2012) demonstrates that income follows a double power law; and Toda and Walsh (2015, 2017) examine the implications of the double power law distribution of consumption growth on asset pricing models.

normal with the inverse-gamma distribution for variance. As such a model is unlikely to accurately capture the distribution of volatility in the data, the subsequent empirical analysis examines a conditional normal model with a nonparametric volatility distribution.

Using more than 925 million distinct observations of stock prices at 1 second intervals, this study investigates whether these model can explain the cross-sectional variation in power law coefficients for the 41 stocks that were included in the Dow Jones Industrial Index at some point from 2003 to 2014. The analysis reveals that the conditional normal model with nonparametric volatility provides a strong fit. Specifically, focusing on the cross-section of stocks, one finds that the power law coefficients implied by this model comove one-for-one with those obtained from return data. More strikingly, examining the stocks in the sample individually one finds evidence that, for the bulk of these stocks, one would not reject the hypothesis that the power law coefficient obtained from the conditional normal model with nonparametric volatility differs from that obtained from return regressions.

The empirical analysis begins by estimating power law coefficients using 15 minute returns from 2003 to 2014 for the 41 stocks in the sample. The results indicate that a power law fits the tail distribution of stocks returns. The estimated power law coefficients are centered around 3, consistent with the findings of Plerou, Gopikrishnan, Amaral, Meyer, and Stanley (1999). The cross-sectional distribution of the estimated power coefficients, which range from 2.09 to 3.46, provides the key variation that the subsequent analysis aims to explain.

Estimating the conditional normal models with time-varying volatility requires measuring volatility at the high frequencies used to measure returns. This poses a challenge, as the realized volatility method developed by Andersen, Bollerslev, Diebold, and Labys (2003) and Barndorff-Nielsen and Shephard (2002) runs into difficulty with market microstructure noise when applied at such high frequencies.<sup>2</sup> As such, the two-scales realized volatility method developed by Zhang, Mykland, and Aït-Sahalia (2005) is used to obtain return volatility within 15 minute intervals, the primary time interval used in the study. Applying this method to returns at 30, 45, or 60 second intervals offset at intervals of 1 second each, one obtains robust measures of return volatility at 15 minute intervals.

First, the study examines a conditional normal model with an inverse-gamma distribution for 15-minute return volatility; this model has the parametric property that the tail

---

<sup>2</sup>Barndorff-Nielsen and Shephard (2004) extends this method to estimate realized covariances and regression coefficients.

distribution of returns is follows a power law with a power coefficient equal to twice the shape coefficient of the inverse-gamma distribution for volatility. Comparing the power law coefficient obtained from 15-minute returns with that obtained by estimating the model reveals a positive relationship between these two sets of coefficients. However, one rejects the hypothesis that the two sets of coefficients are equal.

Next, the study examines a conditional normal model with a nonparametric volatility distribution. For each stock, the power law coefficient implied by this model is obtained from a simulated data set constructed using the empirical distribution for return volatility at 15 minute intervals. A comparison of the model-implied power law coefficients and the corresponding estimates from the tail distribution of return data reveals a close relationship. The correlation coefficient between these two sets of power law coefficients exceeds 0.95. A regression through the origin of the data coefficients on the model-implied coefficients yields slope coefficients between 1.00 and 1.01. These coefficients are statistically indistinguishable from one, providing support for the hypothesis that the two sets of power law coefficients are identical. Testing the model one stock at a time, one finds that the model-implied coefficient falls within the 95 percent confidence interval of the power law coefficient estimated from returns data for the bulk of the stocks in the sample. The ability of the model to fit the data, both in the cross-section and at the individual stock level, indicates that the conditional normal model with time-varying volatility can generate the power law property of stock returns.

One limitation of the above model is that it abstracts from the clustering of volatility observed in the data. As the model-implied power law coefficient is determined by the unconditional distribution of volatility, in principle, this may not be a concern if there are sufficient observations to accurately capture the unconditional distribution of volatility, regardless of the underlying conditional distribution. That said, in order to address this concern, the nonparametric model is estimated using a block sampling scheme that preserves the observed time-series dependency of volatility. One obtains the same key findings from this setting: the slope coefficients of the regression through the origin of the power law coefficients from returns on the model-implied coefficients equal one and, for most of the stocks in the sample taken individually, the model-implied coefficient falls within the 95 percent confidence interval for the coefficient from the return data.

Taken together, these findings suggests that time-varying volatility, a key property of

financial markets emphasized in influential studies such as Engle (1982), Bollerslev (1986) and Schwert (1989), can explain the power law property of high frequency stock returns. The results indicate that economic forces such as time-varying economic uncertainty (see Bloom (2009)) or changes in the arrival rate of information (see Mandelbrot and Taylor (1967) and Kyle and Obizhaeva (2016)) may account for this remarkable regularity of the tail distribution of stock returns.

This study is organized as follows. Section 2 describes the statistical models used in the study; Section 3 details the data used to test these models; Section 4 presents the results of the empirical analysis; Section 5 examines whether the model can fit the entire distribution of returns; and Section 6 concludes.

## 2 Models

This section provides a brief overview of power laws in the tail distributions and presents the main parametric and nonparametric models analyzed in the study.<sup>3</sup>

### 2.1 Power law property

A random variable  $X$  exhibits the power law property in the tail if there exists constants  $A$  and  $\nu$  such that:

$$\lim_{x \rightarrow \infty} x^\nu P(X > x) = A, \tag{1}$$

where  $\nu$  equals the power coefficient. Formally speaking,  $X$  exhibits a Paretian tail.<sup>4</sup> This relationship is typically tested using a log-log plot of the tail distribution  $P(X > x)$  on  $x$ . A linear regression of  $\log(P(X > x))$  on  $\log(x)$  provides an estimate of the power law coefficient,  $\nu$ . As the focus is on the tail distribution of a series, this regression is run only on values of  $x$  that exceed a given threshold.

---

<sup>3</sup>Without loss of generality, we will assume that high-frequency stock returns have zero mean.

<sup>4</sup>While the phrase ‘Paretian tail’ is a more precise description of such distributions, I will follow the bulk of the literature on the power law property of stock returns and refer to this property using the phrase ‘power law’ when discussing it in the context of stock returns.

## 2.2 Parametric model

Consider the following highly stylized parametric distribution for high frequency log returns:

$$\begin{aligned} r_t &\sim N(0, \sigma_t^2), \\ \sigma_t^2 &\sim \text{Inv-Gamma}(\alpha, \beta). \end{aligned} \tag{2}$$

That is, log equity returns follow a conditional normal distribution, with variance ( $\sigma_t^2$ ) drawn from an inverse-gamma distribution with shape parameter  $\alpha$  and scale parameter  $\beta$ . The conditional normal distribution has a long history in understanding asset returns dating back to Engle (1982).<sup>5</sup> The inverse-gamma distribution for variance provides a parsimonious specification that captures time-variation in return volatility.<sup>6</sup>

The joint density function of the above model can be written as:

$$p(r_t, \sigma_t^2) = \frac{1}{\sqrt{2\pi}\sigma_t} \exp\left(\frac{-r_t^2}{2\sigma_t^2}\right) \times \frac{\beta^\alpha}{\Gamma(\alpha)} (\sigma_t^2)^{-\alpha-1} \exp\left(\frac{-\beta}{\sigma_t^2}\right), \tag{3}$$

where  $\Gamma(\cdot)$  denotes the gamma function. Following the derivation in Poirier (1995), one can show that the marginal distribution of returns follows a Student's  $t$ -distribution:

$$p(r_t) = \frac{\Gamma((2\alpha + 1)/2)}{\Gamma(\alpha)\sqrt{(2\pi\beta)}} \left(1 + \frac{r_t^2}{2\beta}\right)^{-(2\alpha+1)/2}. \tag{4}$$

The above  $t$ -distribution has a degree-of-freedom equal to twice the shape parameter of the inverse-gamma distribution,  $2\alpha$ , a mean of 0, and, for  $\alpha > 1$ , a variance equal to  $\frac{\beta}{\alpha-1}$ . Using the fact that the tail distribution of a  $t$ -distribution follows a power law with power coefficient equal to the degree-of-freedom of the  $t$ -distribution, one can write the unconditional tail distribution of returns as:

$$P(r_t > r) \propto r^{-2\alpha}. \tag{5}$$

The above derivation shows that the parsimonious conditional normal distribution shown

---

<sup>5</sup>Andersen, Bollerslev, Diebold, and Ebens (2001) provide evidence that daily stock returns are well described by a conditional normal distribution.

<sup>6</sup>The above specification is also used in the Bayesian literature as a conjugate informative prior.

in Equation (2) implies a power law for the tail distribution of returns.<sup>7</sup> Strikingly, the power law coefficient is inherited directly from the unconditional distribution of the variance of returns, providing a simple testable prediction.

### 2.3 Generalization to models with fat tails for volatility

Other mixture distributions that combine thin-tailed distributions for conditional returns with fat-tailed distributions for volatility may also yield the power law property of the tail distribution. In particular, the following theorem establishes that one can generalize the above derivation of a power law in the tail distribution to a broad class of mixture distributions.<sup>8</sup>

**Theorem 1** *Consider a mixture distribution that consists of a thin-tailed distribution for conditional returns,  $r_t$ , combined with a distribution for volatility,  $\sigma_t$ , that exhibits a Paretian tail (power law) with power coefficient  $\lambda$ . The tail distribution of the unconditional returns obtained by this mixture distribution exhibits a Paretian tail (power law) with power coefficient  $\lambda$ .*

**Proof.** See Appendix A. ■

The empirical analysis of the above proposition focuses on a conditional normal distribution for returns with a nonparametric volatility distribution. I.e., this study also examines the following model:

$$\begin{aligned} r_t &\sim N(0, \sigma_t^2), \\ \sigma_t^2 &\sim f(\cdot), \end{aligned} \tag{6}$$

where  $f(\cdot)$  denotes a nonparametric distribution for variance. The nonparametric specification has the flexibility to match the observed empirical distribution of return volatility. This

---

<sup>7</sup>Weitzman (2007) shows that a  $t$ -distribution for consumption growth arising from parameter uncertainty may help explain the equity premium.

<sup>8</sup>I thank the referee for pointing out this result and outlining the proof.



model is tested by comparing the power law coefficient for the tail distribution of returns obtained from simulating this model using the empirical distribution for return volatility at high frequencies with the corresponding power law coefficient obtained using data on returns.

One limitation of the power law literature in general, and the models (2) and (6) in particular, is that they are not invariant to time-scales, as these distributions are not invariant under addition. In practice, this appears to have little impact, as Plerou, Gopikrishnan, Amaral, Meyer, and Stanley (1999) find that power law coefficients remain mostly constant for returns measured over time intervals ranging from 5 minutes to 1 day.

### 3 Data

The data used in the study are obtained from the Consolidated Trades segment of the Trades and Quotes (TAQ) database accessed through Wharton Research Data Services. The sample comprises all stocks that were included in the Dow Jones Industrial Average at some point from 01/01/2003 to 12/31/2014. Including additions and deletions to the index, this list comprises 41 stocks. These stocks were selected as they are among the most liquid stocks traded in the US equity markets. As such, market microstructure concerns are likely to have a smaller impact on measured returns and volatilities for these stocks.

For each of the stocks in the sample, stock prices are measured at 1 second intervals daily from 01/01/2003 to 12/31/2014. For stocks that had multiple trades within the same second, the stock price is obtained from the median of the reported transaction prices. The sample excludes dates with an early close to trading activity.<sup>9</sup> In order to avoid the effect of information released at the beginning or the end of the trading day, stock prices are measured from 9:45am to 3:45pm. Even taking into account that not all stocks are traded every second, the data set consists of more than 925 million distinct stock price observations.

Following the convention in the realized volatility literature, throughout the study, returns are measured as logarithmic returns, rather than arithmetic returns.<sup>10</sup> Given the short time period over which returns are measured, the logarithmic returns are almost identical to arithmetic returns, and this choice has little effect on the results (see also Kelly (2014)).

---

<sup>9</sup>I thank Yesol Huh for providing data on early closing dates.

<sup>10</sup>Given a stock price process,  $Y_t$ , logarithmic returns are given by  $\log\left(\frac{Y_{t+1}}{Y_t}\right)$ , while arithmetic returns are given by  $\frac{Y_{t+1}}{Y_t} - 1$ .

The empirical analysis in the study primarily uses stock returns and volatilities measured over 15 minute intervals, the time interval used in Gabaix, Gopikrishnan, Plerou, and Stanley (2006).<sup>11</sup> A 15 minute interval provides a sufficient number of price observations to measure volatility within that interval—see the discussion below—while also providing a large number of return observations that enable a precise estimate of the power law coefficient.

### 3.1 Measurement of volatility

The analysis requires measurement of stock return volatility within 15 minute intervals. This poses a challenge, as the traditional approach in the realized volatility literature measures volatility at daily frequencies using either 1- or 5-minute returns (see Andersen, Bollerslev, Diebold, and Labys (2003)). There are too few observations of 1- or 5-minute returns to derive a robust measure of volatility at 15 minute intervals. One cannot address this issue by measuring returns at much higher frequencies, as that increases the market microstructure component of observed returns, thereby adding noise to the realized volatility measures.<sup>12</sup>

This study tackles this challenge by measuring volatility at 15 minute frequencies using the two-scales realized volatility (hereafter, TSRV) method developed by Zhang, Mykland, and Aït-Sahalia (2005) and Aït-Sahalia, Mykland, and Zhang (2011). They note that measuring realized volatility using returns over 5 minute intervals involves a considerable loss of information and develop a method to measure realized volatility utilizing stock price data at much higher frequencies. This method involves first computing realized volatility using returns at a frequency typically used in the literature, such as a 1 minute frequency. Based on the observation that there are many such overlapping 1-minute return intervals within a period of interest, the above studies show that the average of the realized volatility measure over these overlapping intervals, adjusted for a bias correction term involving realized volatility calculated using returns at the higher frequency, provides an estimate of realized volatility over the desired interval.<sup>13</sup> For instance, using a smaller time scale of 1 second, one could obtain realized volatility based on the average of 60 different 1-minute returns, each

---

<sup>11</sup>The stock price at the end of each 15 minute interval is obtained as the last price at which the stock traded prior to the end of the 15 minute interval.

<sup>12</sup>Hansen and Lunde (2006) discusses the effect of market microstructure noise on realized volatility measures as the interval over which returns are measured shrinks.

<sup>13</sup>The TSRV measure used in the study also incorporates the small sample refinement proposed in Aït-Sahalia, Mykland, and Zhang (2011).

offset by 1 second. This method enables one to measure realized volatility using stock price information at the much higher frequency by which the returns are offset rather than the lower frequency at which the squared returns are measured. The Appendix provides further details on the measurement of volatility using this method.

The results reported in this study are obtained using the TSRV method applied at frequencies of 30, 45 and 60 seconds offset at 1 second intervals. This enables one to use up to 900 observations of stock prices at 1 second intervals to measure realized volatility within each 15 minute interval. Observations for which the bias correction term causes one of the TSRV measures to be negative are discarded from the sample. Finally, the realized volatility measures in the sample are Winsorized at the 0.01th percentile to remove the effect of extreme outliers.<sup>14</sup> This affects only a very small fraction (0.02%) of the observations, and is done primarily to limit the effect of outliers on the parameter estimates of the inverse-gamma model.

### 3.2 Summary statistics

Table 1 reports summary statistics for the 15-minute return and volatility data used in the subsequent analysis. Panel A reports univariate statistics; Panel B reports the correlation coefficients among the TSRV volatility measures obtained using different return frequencies; and Panel C reports the correlation coefficients among the corresponding realized volatility measures. Taken together, the data set used in the analysis includes more than 2 million observations of stock returns and volatilities measured over 15-minute intervals (note that many more stock price observations were used to construct these volatility measures). On a per stock basis, the data set contains about 70,000 15-minute return and volatility data for most of the 41 stocks in the sample. The large sample size helps ensure that small sample concerns do not affect the estimated power law coefficients and that the data set captures the unconditional distributions of the return volatility of each stock.

The univariate statistics indicate that the mean 15-minute return among the stocks in the sample is slightly positive, reflecting the fact that overall stock returns were positive over the sample period. The corresponding median return equals zero. The statistics also indicate that the TSRV measures provide a robust estimate of realized volatility at 15 minute

---

<sup>14</sup>That is, observations with realized volatility above the 99.99th percentile are set to the value at this percentile while observations below the 0.01th percentile are set to that value.

intervals over the three return frequencies used in the study. Moving from measuring squared returns at a 60 second interval to a 30 second interval leads to only a slight increase in the mean and dispersion of realized volatilities, suggesting that the TSRV approach manages to successfully account for market microstructure noise in the data while incorporating a sizeable quantity of stock price information. The robustness of this approach is further supported by the correlation matrix shown in Panel B, which indicates that the three TSRV measures have a very high correlation with one another.

Panel C reports the corresponding correlation matrix for realized volatility measures at 15-minute intervals obtained using the approach in Andersen, Bollerslev, Diebold, and Labys (2003) applied to squared returns at 30, 45 and 60 second intervals. The results indicate that while the realized volatility measures are closely correlated, the correlations among them are not as strong as those for the TSRV volatility measures. This suggests that the TSRV approach provides a more robust measure of volatility within 15 minute intervals than the realized volatility method.

## 4 Results

The empirical analysis tests the models presented in Section 2 by comparing the cross-section of power law coefficients obtained from the tail distribution of returns with the corresponding model-implied coefficients. That is, the tests examine whether models with time-varying volatility can explain the observed variation in the power law coefficients across the stocks in the sample. In comparison, prior studies such as Gabaix, Gopikrishnan, Plerou, and Stanley (2006) have focused on explaining a single power law coefficient obtained by aggregating information from different stocks.

### 4.1 Power law coefficients from stock returns

The power law coefficients for the tail distribution of stock returns are obtained from regressions of the log probability that the absolute stock return exceeds a given value on the log of the absolute return. These regressions are carried out for absolute stock returns greater than the 95<sup>th</sup> percentile. The standard error of each regression coefficient is obtained using the method in Gabaix and Ibragimov (2011). As the sample consists of about 70,000 returns for

most stocks, each regression uses about 3,500 observations on tail probabilities, mitigating small sample concerns.

Figure 1 presents log-log plots of the tail distribution of the absolute value of 15-minute returns for each of the stocks in the sample. It also shows the fitted line obtained from a regression of the log tail probability on the absolute stock return. As the figure shows, a linear regression fits the tail distribution for the absolute returns of the stocks used in the sample well, indicating that the tail distribution of returns follows a power law, consistent with the literature. For some of the stocks in the sample, the fitted line exceeds the implied probability that the absolute return will be greater than a given value for the small number of observations near the very tail of the distribution.<sup>15</sup>

Figure 2 gathers the above results and plots the distribution of the estimated power law coefficients for all the stocks in the sample. As the figure indicates, the distribution of the estimated power law coefficients is centered around 3, with a mean value of 2.92 and a median of 3.00. This result matches the finding in Gabaix, Gopikrishnan, Plerou, and Stanley (2003) and others that high frequency stock returns exhibit a power law coefficient of around 3. The estimated power law coefficients exhibit some dispersion, with a cross-sectional standard deviation of 0.32. The subsequent empirical analysis examines whether the models presented in Section 2 can capture this cross-sectional variation. In comparison, prior studies have not focused on examining the cross-section of power law coefficients observed in the data.

## 4.2 Estimation of parametric model

As noted in Section 2.2, a conditional normal model with an inverse-gamma volatility distribution implies that the tail distribution of stock returns follows a power law. This section reports the results from testing this model by comparing the power law coefficients obtained from the data with the model-implied coefficients, given by 2 times the shape coefficient of the inverse-gamma distribution for stock return variance measured at 15 minute intervals. The shape and scale parameters of the inverse-gamma distributions are obtained using maximum likelihood estimation.

---

<sup>15</sup>This effect is visually enhanced by the log-log plot, which gives equal prominence to values that have tail probabilities between  $10^{-4}$  and  $10^{-5}$  as those that have a tail probability between  $10^{-2}$  and  $10^{-3}$ , even though there are far fewer observations of the former.

Figure 3 plots the power law coefficient obtained from the regression analysis in Section 4.1 on two times the shape parameters of the inverse-gamma distributions for return variance. Panels A, B and C, respectively, plot this correspondence for the variance distributions obtained using the TSRV method with volatility measured within each 15-minute window using returns at 30, 45 and 60 second intervals offset at 1 second. The figures also report the slope of a regression through the origin of the power law coefficients from the data on the model-implied coefficients.

The null hypothesis in this (and subsequent related) analysis is whether the power law coefficients implied by the model equal their data counterparts. This section examines the null hypothesis by testing whether the slope coefficient of a regression through the origin of the power law coefficients from the data on the model-implied coefficients equals 1. Notably, the constant term in the regression equals zero by construction to reflect the fact that a simple linear model with a non-zero constant term would be a misspecified model for examining whether the two sets of coefficients are identical. Imposing the restriction that the constant term in the regression equals zero implies that the fitted line obtained from the regression shown in Figure 3 (and in subsequent figures) does not line up with a more conventional best-fit line obtained from a linear regression with a non-zero constant.

As the figure indicates, there exists a clear positive relationship between the power law coefficients obtained from stock returns and the power law coefficient implied by the conditional normal model given in Equation (2). The slope of the linear regressions through the origin for the three volatility measures are positive and significant, albeit significantly greater than 1, implying that one rejects the null that the two sets of coefficients are identical. While stocks with high/(low) model-implied power law coefficients also have high/(low) coefficients in the data, there is a greater dispersion in the model-implied coefficients than in the data. Further, the model-implied coefficients are lower, on average, than the coefficients from the returns data.

The estimated model uses a highly stylized distribution for volatility. Comparing the empirical distribution for variance to that obtained by fitting the inverse-gamma distribution, one finds that the inverse-gamma distribution has greater mass at the center of the distribution and less mass at the tails than in the data. We next examine whether a conditional normal model with a more general, nonparametric volatility distribution can provide a better fit for the power law property of stock returns.

### 4.3 Estimation of conditional normal model with nonparametric volatility

This section examines whether the conditional normal model with a nonparametric distribution for volatility shown in Equation (6) can generate the power law coefficients observed in the data. Using a nonparametric specification allows one to match the observed volatility distribution in the data. However, this poses the limitation that one no longer obtains a parametric derivation of the power law coefficient implied by the model. As such, the model-implied power law coefficients are obtained using simulation methods.

Specifically, for each stock, one obtains the model-implied power law coefficient from a simulated data set of returns,  $r_i, i = 1, \dots, S$  constructed as follows:

1. Obtain a draw for variance,  $\sigma_i^2$ , by sampling with replacement from the empirical distribution for variance,  $f(\cdot)$ .
2. Obtain a draw for returns,  $r_i$ , by sampling from a normal distribution with mean zero and variance  $\sigma_i^2$ .

This simulated sample has a size,  $S$ , of 100 times the number of 15-minute return observations in the data, enabling a precise estimate of the power law coefficient implied by the model; the standard errors of the power law coefficients obtained from the simulated data are all less than 0.01. As such, the subsequent analysis abstracts from the estimation error for the model-implied power law coefficients.<sup>16</sup>

#### 4.3.1 Cross-sectional evidence

Figure 4A plots the power law coefficients obtained from regression analyses of 15-minute returns data on the power law coefficients obtained using simulated data from the conditional normal model with nonparametric volatility distribution. The volatility distribution underlying the model is given by the empirical distribution of 15-minute return volatility obtained using the TSRV method with 30-second returns offset at 1-second intervals. Figures 4B and 4C, respectively, plot the corresponding results for the simulated data sets from the

---

<sup>16</sup>A Monte Carlo analysis using 400 different simulation samples confirms that there is little simulation error in the model-implied coefficients and that this simulation error has no impact on the subsequent analysis.

conditional normal models with volatility obtained using the TSRV method applied to 45- and 60-second returns offset at 1-second intervals. The figures also present the slope coefficients, associated 95 percent confidence intervals and fitted lines obtained from a regression through the origin of the power law coefficients from the returns data on the model-implied power law coefficients. As noted earlier, imposing the restriction that the constant term of the regression equals zero enables one to test the hypothesis that the two sets of power law coefficients are identical.

The figures indicate a striking concordance between the power law coefficients obtained from 15-minute returns and those obtained from the conditional normal models. For each of the three 15-minute return volatility measures, stocks with high/(low) model-implied power law coefficients also have high/(low) power law coefficients in the return data. Indeed, the correlation coefficient between the two sets of power law coefficients ranges from 0.96 to 0.98 across the three volatility measures. In addition, the dispersion of the model-implied power law coefficients is also quite similar to that of the data coefficients.

Turning to the regression analysis, one finds that the slope of a regression line through the origin of the power law coefficient from 15-minute returns on the corresponding power law coefficient from the conditional normal models ranges from 1.004 to 1.007. For each of the three volatility measures, the 95 percent confidence interval of the estimated slope coefficient encompasses one. The corresponding  $p$ -values for the tests that the slope coefficients reported in Figures 4A, 4B and 4C equal one are 0.41, 0.27, and 0.06, respectively. This indicates that, for all these volatility measures, one would not reject the hypothesis that the two sets of power law coefficients are identical. This close fit indicates that the model can explain the cross-section of power law coefficients observed in the data.

Note that the above finding is a stronger result than that typically obtained in empirical studies, which focus on examining whether a given coefficient is positive or not. Instead, the analysis tests the hypothesis that the two sets of coefficients are identical by examining whether the slope coefficient equals a specific value, one, and finds that one fails to reject this hypothesis at conventional significance levels.

What explains the improved performance of the nonparametric model compared to the results obtained using the parametric model? Examining the fit of the inverse-gamma distribution for volatility, one finds that the model implies less frequent realizations for inverse volatility than observed in the data around the mode of the distributions and more frequent



realizations for very low values. As such, the inverse-gamma distribution predicts a higher frequency of very high volatility realizations than observed in the data. As high volatility realizations are associated with high absolute returns, the return distributions implied by the parametric model has fatter tails than the corresponding distributions implied by the nonparametric model. The fatter tails in the parametric model leads to lower estimates for the power law coefficients  $\nu$ , as can be seen by comparing the results reported in Figure 3 with those reported in Figures 4A, 4B and 4C.

The above confidence intervals and  $p$ -values are obtained using OLS standard errors applied to a  $t$ -distribution with 40 degrees of freedom. The analysis does not use heteroskedasticity robust standard errors due to the fact that the finite sample properties of the resulting  $t$ -statistics are not well known. As such, it is necessary to examine whether the above regressions suffer from heteroskedasticity in the error term. Applying the White (1980) and Breusch and Pagan (1979) heteroskedasticity tests to the regressions for the three volatility measures, one finds little significant evidence of heteroskedasticity.<sup>17</sup> This finding is confirmed by a visual examination of the residuals. Consistent with these findings, the heteroskedasticity robust standard errors for the slope coefficients are within 2 basis points of the OLS standard errors.

### 4.3.2 Evidence at the individual stock level

Another way of evaluating the model is to investigate whether, for a given stock, the power law coefficient obtained from the simulated data falls within the 95 percent confidence interval of the power law coefficient estimated using 15-minute returns data. This examines whether the model can explain the data at the individual stock level, thus providing a sterner test of the model. Table 2 presents the results of this analysis. For each of the stocks in the sample, it reports the power law coefficient estimated from the returns data, the associated 95 percent confidence interval, and the model-implied power law coefficients obtained using simulated data for each of the three measures of 15-minute return volatility used in the study. \* and \*\*, respectively, denote stocks for which one can reject the hypothesis the model-implied power law coefficient equals that obtained from the return regression at the 95 and 99 percent significance levels.

---

<sup>17</sup>The  $p$ -values for the White test equals 0.04, 0.21 and 0.45, respectively, while the  $p$ -values for the Breusch-Pagan test range from 0.34 to 0.52.

The results show that the model-implied power law coefficient falls within the 95 percent confidence interval for the regression estimate from returns data for between 36 to 38 of the 41 stocks in the sample. This indicates that, for most of the stocks in the sample taken individually, one would not reject the hypothesis that the power law coefficient implied by the conditional normal model is different from the corresponding coefficient obtained from the regression analysis using 15-minute returns. The ability of the model to fit the power law coefficients obtained from returns at the individual stock level provides clear evidence that the conditional normal model with time-varying volatility can explain the power law property of stock returns.

### 4.3.3 Volatility clustering

The conditional normal models used in the study abstract from the time-series dependency of volatility, highlighted in Engle (1982), Bollerslev (1986) and Bollerslev and Wright (2001), among others. I.e., the estimation of the non-parametric model employs random sampling of the empirical volatility distribution. This approach is used as the models' implications are based on the unconditional volatility distribution and, in principle, one would be able to accurately obtain the model-implied power law coefficient if one had sufficient observations to capture the unconditional distribution for volatility, regardless of the underlying conditional distribution. The empirical distribution includes about 70,000 15-minute volatility observations over 10 years for most of the stocks in the sample, suggesting that one has sufficient observations to capture the unconditional distribution. Nonetheless, this section uses an alternate sampling scheme to examine whether the results differ when one estimates the model while accounting for the time-series dependency of the volatility distribution.

The nonparametric conditional normal model with volatility clustering is estimated using a block sampling method. That is, the model-implied volatility distribution is obtained by sampling from blocks of 120 consecutive observations for 15-minute volatility from the corresponding empirical distribution.<sup>18</sup> This sampling approach, akin to the block bootstrap method, helps preserve the time series correlation of volatility in the data. The model-implied power law coefficient is obtained from a simulated sample of returns obtained from a conditional normal distribution with these volatility draws.

---

<sup>18</sup>The overall simulation sample size is approximately equal to 100 times the number of 15-minute return observations, similar to the simulation sample size in the estimation without volatility clustering.

Figure 5 presents scatter plots of the power law coefficients obtained from the return regression against the coefficients implied by the nonparametric model with volatility clustering. Panels A, B, and C, respectively, present results from measuring volatility at 15 minute intervals using the TSRV method applied to 30, 45, and 60 second intervals offset at one second. Across the three volatility measures and the 41 stocks in the sample, the volatility distribution in the simulated data exhibit an auto-correlation of 0.581, close to the mean auto-correlation coefficient in the volatility data of 0.588. This indicates that the sampling method captures the time-series dependency of volatility in the data.

As the figure indicates, one continues to obtain the key result using this specification. That is, the slope coefficient of the regression through the origin of the power law coefficient from the return regression on the model-implied coefficient is statistically and economically indistinguishable from one for each of the three volatility measures. The 95 percent confidence intervals for the slope coefficient encompass one, and the  $p$ -values for testing whether the slope coefficient equals one ranges from 0.11 to 0.45. This indicates that, for all the three volatility measures, one would not reject the hypothesis that the model-implied power law coefficients differ from those obtained from the return regressions.

Furthermore, comparing the results at the individual stock level one finds that for the three volatility measures, the model-implied power law coefficient falls within the 95 percent confidence interval for the coefficient from the return regression for between 36 to 38 stocks in the sample, similar to that obtained from the nonparametric model without volatility clustering.

One limitation of the analysis is that it does not account for the leverage effect of returns on future volatility (see Christie (1982)). This partly reflects the challenge of nonparametric sampling from the unconditional distribution of volatility while incorporating a leverage effect. It also reflects the fact that, empirically, lagged volatility has much stronger predictive power for future volatility than lagged returns, which suggests that incorporating a leverage effect is unlikely to substantially change the above results.

#### 4.3.4 Additional results

The previous results were obtained using a 15 minute time interval to measure stock returns and volatility as this time interval has been used in prior related studies, and as it enabled sufficient data on prices at a one-second interval to measure return volatility using the TSRV

method. Examining whether the results are robust to using 10 minutes and 30 minutes to measure returns and volatility, one obtains similar results.<sup>19</sup>

## 5 Distribution of normalized returns

One possible concern is that, by focusing exclusively on understanding the power law properties of returns, the above analysis lacks power to reject the model. This section examines whether the conditional normal model with non-parametric volatility can fit the entire distribution of stock returns, thus providing a more powerful test of the model. The analysis focuses on the observation that the conditional normal model implies that returns normalized by volatility should follow a normal distribution with variance equal to 1. I.e.,

$$\frac{r_t}{\sigma_t} \sim N(0, 1).$$

Table 3 presents the mean, standard deviation, skewness and kurtosis of the normalized returns for all the stocks in the sample, pooled together. The normalized returns have a slightly positive mean, consistent with the mean return reported in Table 1. Notably, one finds that the standard deviation of the normalized returns is less than one, indicating that normalized returns are less volatile than implied by the model. Indeed, a Kolmogorov-Smyrnoff test for normality rejects the hypothesis that the normalized returns follow a normal distribution for almost all the stocks in the sample, taken individually. Finally, while the normalized returns exhibit very little skewness, their kurtosis for two of the three volatility measures differs somewhat from the value of 3 that would be obtained from a standard normal.

One can further understand the distribution of normalized returns by examining the joint probability density of volatility (sorted into ten deciles) and normalized returns for the pooled sample, shown in Figure 6. This joint distribution is estimated using a kernel estimator. Panel A presents the joint density for normalized returns and volatility for ten volatility deciles. Panel B presents the corresponding contour plots for the density of normalized returns. These figures are analogous to Figure 1 of Ioannides and Overman (2003), who

---

<sup>19</sup>These results are presented in the online Appendix.

present similar density and contour plots for city size and growth rates.<sup>20</sup>

The conditional normal model implies that the distribution of normalized returns would be independent of volatility. As seen in Panel A, while the density functions for normalized returns look similar across the volatility deciles, there is a clearly observable difference near zero. This occurs due to the fact that the data contains a non-negligible number of observations where the 15-minute return equals exactly zero. This bunching reflects the decimal pricing of individual stocks, which implies that the minimum price change is one cent.<sup>21</sup> As such, there are time intervals in the data where, even though there are intermediate trades within a 15 minute interval, the overall return over this period equals exactly zeros.<sup>22</sup> When the volatility of a return is low, the likelihood of such a zero return increases, leading to a higher density of such observations at the lowest volatility decile, thus resulting in the more pronounced ridge observed in Panel A. This pattern can also be seen clearly in the contour plot shown in Panel B, where the lowest volatility deciles have higher peak densities, and thus more contour lines, than the higher volatility deciles. That said, the bunching at zero occurs even for the high volatility deciles, leading to the rejection of the hypothesis that the conditional distribution of normalized returns is normal for all volatility deciles.

The fact that the market microstructure reasons lead to a bunching of returns at zero, thus resulting in a non-normal distribution for normalized returns does not invalidate the key findings of Section 4 regarding the tail distribution of returns. Namely, that models with time-varying volatility can help explain the cross-sectional distribution of power law coefficients, and that, among these models, the conditional normal with nonparametric volatility provides a strong fit.

## 6 Conclusion

This study examines whether models with time-varying volatility can help explain the observed tail distribution of stock returns. Examining a simple parametric model with an inverse-gamma distribution for variance, one finds a positive relationship between the model-

---

<sup>20</sup>See also Figure 5 of Eeckhout (2004).

<sup>21</sup>Strictly speaking, these observations should result in an infinite density at zero due to their non-zero point mass. The kernel estimator smooths through that, resulting in the observed ridge pattern.

<sup>22</sup>I.e., the fact that stock prices are rounded to one cent lead to observations where the observed price change equals zero while the underlying price process may exhibit a very small change.

implied power law coefficients and the power law coefficients obtained from the data; however, the model-implied coefficients exhibit greater dispersion than the data coefficients. The investigation of a conditional normal model with a nonparametric distribution for volatility reveals a close relationship between the model-implied power law coefficients and the corresponding data values. A cross-sectional analysis using a regression through the origin yields evidence that the two sets of power law coefficients are identical. Furthermore, comparing the model-implied power law coefficients with the corresponding coefficients from the data at the individual stock level, one fails to reject the hypothesis that the model-implied coefficient equals the data coefficient for most of the stocks in the sample, providing evidence in favor of the model at the individual stock level. These results are generally robust to different time intervals for measuring volatility and to a sampling approach that accounts for volatility clustering. Finally, one finds that the distribution of normalized returns given volatility is fairly similar across volatility deciles, except for a bunching at zero that is more pronounced at lower volatilities.

These findings indicate that a conditional normal model with time-varying volatility can explain the power law property of stock returns. Economically, this suggests that either changes in uncertainty over time or changes in the arrival rate of information may explain the tail distribution of stock returns. Econometrically, the findings indicate that mixture distributions may help provide insights into why many financial and economic data series exhibit power law properties. Further research into the channels underpinning the observed power law distributions of other series, such as the distribution of firms and city sizes, may prove fruitful.

## References

- Aït-Sahalia, Yacine, Per A. Mykland, and Lan Zhang, 2011, Ultra high frequency volatility estimation with dependent microstructure noise, *Journal of Econometrics* 160, 160–175.
- Andersen, Torben G., Tim Bollerslev, Francis X. Diebold, and Heiko Ebens, 2001, The distribution of realized stock return volatility, *Journal of Financial Economics* 61, 43–76.
- Andersen, Torben G., Tim Bollerslev, Francis X. Diebold, and Paul Labys, 2003, Modeling and forecasting realized volatility, *Econometrica* 71, 579–625.
- Axtell, Robert L., 2001, Zipf distribution of U.S. firm sizes, *Science* 293, 1818–1820.
- Barndorff-Nielsen, Ole E., and Neil Shephard, 2002, Econometric analysis of realized volatility and its use in estimating stochastic volatility models, *Journal of the Royal Statistical Society B* 64, 253–280.
- Barndorff-Nielsen, Ole E., and Neil Shephard, 2004, Econometric analysis of realized covariation: high frequency based covariance, regression, and correlation in financial economics, *Econometrica* 72, 885–925.
- Bloom, Nick, 2009, The impact of uncertainty shocks, *Econometrica* 77, 623–685.
- Bollerslev, Tim, 1986, Generalized autoregressive conditional heteroskedasticity, *Journal of Econometrics* 31, 307–327.
- Bollerslev, Tim, and Jonathan H. Wright, 2001, High-frequency data, frequency domain inference and volatility forecasting, *Review of Economics and Statistics* 83, 596–602.
- Breusch, Trevor, and Adrian R. Pagan, 1979, A simple test for heteroskedasticity and random coefficient variation, *Econometrica* 47, 1287–1294.
- Christie, Andrew A., 1982, The stochastic behavior of common stock variances: Value, leverage and interest rate effects, *Journal of Financial Economics* 10, 407–432.
- Clark, Peter K., 1973, A subordinated stochastic process model with finite variance for speculative prices, *Econometrica* 41, 135–155.
- Eeckhout, Jan, 2004, Gibrat’s law for (all) cities, *American Economic Review* 95, 1429–1451.

- Engle, Robert F., 1982, Autoregressive conditional heteroskedasticity with estimates of the variance of United Kingdom inflation, *Econometrica* 50, 987–1007.
- Gabaix, Xavier, Parameswaran Gopikrishnan, Vasiliki Plerou, and H. Eugene Stanley, 2003, A theory of power-law distributions in financial market fluctuations, *Nature* 423, 267–270.
- Gabaix, Xavier, Parameswaran Gopikrishnan, Vasiliki Plerou, and H. Eugene Stanley, 2006, Institutional investors and stock market volatility, *Quarterly Journal of Economics* 121, 461–504.
- Gabaix, Xavier, and Rustam Ibragimov, 2011, Rank  $-1/2$ : A simple way to improve the OLS estimation of tail exponents, *Journal of Business and Economic Statistics* 29, 255–293.
- Gopikrishnan, Parameswaran, Vasiliki Plerou, Xavier Gabaix, and H. Eugene Stanley, 2000, Statistical properties of share volume traded in financial markets, *Physical Review E* 62, 4493–4496.
- Hansen, Peter R., and Asger Lunde, 2006, Realized variance and market microstructure noise, *Journal of Business Economics and Statistics* 24, 127–161.
- Ioannides, Yannis M., and Henry G. Overman, 2003, Zipf’s law for cities: An empirical investigation, *Regional Science and Urban Economics* 33, 127–137.
- Kelly, Bryan, 2014, The dynamic power law model, *Extremes* 17, 557–583.
- Kyle, Albert S., and Anna A. Obizhaeva, 2016, Market microstructure invariance: Empirical hypotheses, *Econometrica* 84, 1345–1404.
- Mandelbrot, Benoit, and Howard M. Taylor, 1967, On the distribution of stock price differences, *Operations Research* 15, 1057–1062.
- Plerou, Vasiliki, Parameswaran Gopikrishnan, Luis A. Nunes Amaral, Martin Meyer, and H. Eugene Stanley, 1999, Scaling of the distribution of price fluctuations of individual companies, *Physical Review E* 60, 6519–6529.
- Poirier, Dale J., 1995, *Intermediate Statistics and Econometrics*. (MIT Press Cambridge, MA).



- Rozenfeld, Hernan D., Diego Rybski, Xavier Gabaix, and Hernan A Makse, 2011, The area and population of cities: New insights from a different perspective on cities, *American Economic Review* 101, 2205–2225.
- Schwert, William G., 1989, Why does stock market volatility change over time?, *Journal of Finance* 44, 1115–1153.
- Stumpf, Michael P. H., and Mason A. Porter, 2012, Critical truths about power laws, *Science* 335, 665–666.
- Toda, Alexis A., 2012, The double power law in income distribution: Explanations and evidence, *Journal of Economic Behavior and Organization* 84, 364–381.
- Toda, Alexis A., and Kieran Walsh, 2015, The double power law in consumption and implications for testing Euler equations, *Journal of Political Economy* 123, 1177–1200.
- Toda, Alexis A., and Kieran Walsh, 2017, Fat tails and spurious estimations of consumption-based asset pricing models, *Journal of Applied Econometrics*.
- Weitzman, Martin L., 2007, Subjective expectations and asset-return puzzles, *American Economic Review* 97, 1102–1130.
- White, Halbert, 1980, A heteroskedasticity-consistent covariance matrix estimator and a direct test for heteroskedasticity, *Econometrica* 48, 817–838.
- Zhang, Lan, Per A. Mykland, and Yacine Aït-Sahalia, 2005, A tale of two time scales: Determining integrated volatility with noisy high-frequency data, *Journal of the American Statistical Association* 100, 1394–1411.

# Appendix

## A Proof

**Theorem 1** *Consider a mixture distribution that consists of a thin-tailed distribution for conditional returns,  $r_t$ , combined with a distribution for volatility,  $\sigma_t$ , that exhibits a Paretian tail (power law) with power coefficient  $\lambda$ . The tail distribution of the unconditional returns obtained by this mixture distribution exhibits a Paretian tail (power law) with power coefficient  $\lambda$ .*

**Proof.** Let  $h(\cdot)$  denote the density of conditional returns when the standard deviation is normalized to 1. By assumption,  $h(\cdot)$  will be thin-tailed. By a change of variable, the density when the standard deviation equals  $\sigma$  is given by

$$\frac{1}{\sigma} h\left(\frac{r_t}{\sigma}\right).$$

Let  $g(\sigma)$  denote the density for volatility. By assumption,  $g(\sigma)$  exhibits a Paretian tail:

$$\lim_{\sigma \rightarrow \infty} \sigma^{\lambda+1} g(\sigma) = C, \tag{A.1}$$

where  $C$  is a positive constant.

Then, the unconditional density for returns is given by the following:

$$p(r_t) = \int_0^\infty \frac{1}{\sigma} h\left(\frac{r_t}{\sigma}\right) d\sigma. \tag{A.2}$$

Fix any  $\bar{\sigma}$ . One can write the above equation as

$$p(r_t) = \int_0^{\bar{\sigma}} \frac{1}{\sigma} h\left(\frac{r_t}{\sigma}\right) d\sigma + \int_{\bar{\sigma}}^\infty \frac{1}{\sigma} h\left(\frac{r_t}{\sigma}\right) d\sigma. \tag{A.3}$$

Multiply both sides by  $r_t^{\lambda+1}$  and let  $r_t \rightarrow \infty$  to obtain.

$$\lim_{r_t \rightarrow \infty} r_t^{\lambda+1} p(r_t) = \lim_{r_t \rightarrow \infty} \int_0^{\bar{\sigma}} \frac{1}{\sigma} r_t^{\lambda+1} h\left(\frac{r_t}{\sigma}\right) d\sigma + \lim_{r_t \rightarrow \infty} \int_{\bar{\sigma}}^{\infty} \frac{1}{\sigma} r_t^{\lambda+1} h\left(\frac{r_t}{\sigma}\right) d\sigma. \quad (\text{A.4})$$

As  $h(\cdot)$  is thin-tailed, one obtains that

$$\lim_{r_t \rightarrow \infty} r_t^{\lambda+1} h\left(\frac{r_t}{\sigma}\right) = 0.$$

Thus, the first integral vanishes, and we obtain

$$\lim_{r_t \rightarrow \infty} r_t^{\lambda+1} p(r_t) = \lim_{r_t \rightarrow \infty} \int_{\bar{\sigma}}^{\infty} \frac{1}{\sigma} r_t^{\lambda+1} h\left(\frac{r_t}{\sigma}\right) d\sigma. \quad (\text{A.5})$$

Using the change of variables  $x = \frac{r_t}{\sigma}$ , one can write the above integral as

$$\int_{\bar{\sigma}}^{\infty} r_t^{\lambda+1} h\left(\frac{r_t}{\sigma}\right) \frac{d\sigma}{\sigma} = \int_0^{r_t/\bar{\sigma}} h(x) g\left(\frac{r_t}{x}\right) r_t^{\lambda+1} \frac{dx}{x}.$$

Furthermore, the latter integral can be written as

$$\int_0^{r_t/\bar{\sigma}} h(x) g\left(\frac{r_t}{x}\right) r_t^{\lambda+1} \frac{dx}{x} = \int_0^{\infty} x^{\lambda} h(x) g\left(\frac{r_t}{x}\right) \left(\frac{r_t}{x}\right)^{\lambda+1} \iota(x \leq r_t/\bar{\sigma}) dx,$$

where  $\iota(\cdot)$  denotes the indicator function. Plug this into Equation (A.5) to obtain that

$$\lim_{r_t \rightarrow \infty} r_t^{\lambda+1} p(r_t) = \lim_{r_t \rightarrow \infty} \int_0^{\infty} x^{\lambda} h(x) g\left(\frac{r_t}{x}\right) \left(\frac{r_t}{x}\right)^{\lambda+1} \iota(x \leq r/\bar{\sigma}) dx. \quad (\text{A.6})$$

From Equation (A.1), one obtains that

$$\lim_{r_t \rightarrow \infty} g\left(\frac{r_t}{x}\right) \left(\frac{r_t}{x}\right)^{\lambda+1} = C.$$

Furthermore as  $\lim_{r_t \rightarrow \infty} \iota(x \leq r/\bar{\sigma}) = 1$ , by the dominated convergence theorem we can

obtain the simplification that

$$\lim_{r_t \rightarrow \infty} r_t^{\lambda+1} p(r_t) = \int_0^{\infty} x^{\lambda} h(x) C dx. \quad (\text{A.7})$$

As  $h(x)$  denotes a density function with mean 0 and standard deviation 1, the last integral equals a constant,  $\bar{C}$ . I.e.,

$$\lim_{r_t \rightarrow \infty} r_t^{\lambda+1} p(r_t) = \bar{C}, \quad (\text{A.8})$$

implying the desired property that the unconditional distribution of returns,  $r_t$ , exhibits a Paretian tail with power law coefficient  $\lambda$ .

■

## B Measurement of volatility using the TSRV method

This section provides a brief summary of the TSRV method developed in Zhang, Mykland, and Aït-Sahalia (2005) and Aït-Sahalia, Mykland, and Zhang (2011) that is used in the study to measure realized volatility at 15-minute intervals. For further details, see the above studies.

Let  $\tilde{r}_t$  denote the true return process of a security and  $\sigma_{t,15}$  denote the volatility of the true return process over a 15-minute window. Let  $r_{t,15}^1$  denote the observed returns of the security over 1-second intervals within the 15-minute window. In the absence of market microstructure noise, one could estimate the volatility of the true return process based on the realized volatility of observed returns.

$$RV_{t,15}^{\text{all}} = \frac{1}{n} \sum (r_{t,15}^1)^2, \quad (\text{B.1})$$

where the number of 1-second returns within a 15-minute window,  $n$ , equals 900.

The above estimate is biased in the presence of market microstructure noise, leading one to instead estimate realized volatility by measuring returns over a sparse interval, such as a 60-second or a 300-second window. Letting  $r_{t,15}^{60}$  denote returns over 60 second windows, the

corresponding realized volatility estimate is given by

$$RV_{t,15}^{\text{sparse}} = \frac{1}{\bar{n}} \sum (r_{t,15}^{60})^2, \quad (\text{B.2})$$

where  $\bar{n}$  denotes the number of such returns. While this estimate is much less influenced by market microstructure noise, it involves a loss of information.

The method of Aït-Sahalia, Mykland, and Zhang (2011) incorporates additional information into the estimate of realized volatility by measuring returns at 60-second windows offset at  $k$  seconds. Denote the corresponding 60-second returns offset at  $k$  seconds by  $r_{t,15}^{60,k}$ . The realized volatility measure obtained from these returns is given by:

$$RV_{t,15}^{\text{sparse},k} = \frac{1}{\bar{n}} \sum \left( r_{t,15}^{60,k} \right)^2. \quad (\text{B.3})$$

One can construct  $K$  such measures for  $k \in 1, 2, \dots, K$ .

Aït-Sahalia, Mykland, and Zhang (2011) show that one could obtain a consistent estimate of the volatility of the true return process by taking the average of the above realized volatility values and adjusting for a bias-correction term involving the realized volatility measure obtained using 1-second returns. Their TSRV estimator is given by

$$TSRV_{t,15} = \frac{1}{K} \sum_{k=1}^K RV_{t,15}^{\text{sparse},k} - \frac{\bar{n}}{n} RV_{t,15}^{\text{all}}. \quad (\text{B.4})$$

One limitation of the above estimator is that it convergences only at  $n^{1/6}$ . As such, this study measures realized volatility using the following small-sample refinement proposed by Aït-Sahalia, Mykland, and Zhang (2011):

$$TSRV_{t,15}^{\text{adj}} = \left( 1 - \frac{\bar{n}}{n} \right) TSRV_{t,15}. \quad (\text{B.5})$$

The analysis in the study employs the above method to measure volatility at 15 minute intervals. The Online Appendix provides results at the 10 and 30 minute intervals. The time period at which returns are offset,  $k$ , increments at one second. For each time interval, three

estimates of TSRV are obtained using squared returns measured over 30, 45 and 60 second intervals (the corresponding  $K$  values equal 30, 45 and 60, respectively).

Table 1: Data moments

Calculations based on Consolidated Trades data from the TAQ database. The table reports summary statistics for intraday stock returns and return volatility measured at 15 minute intervals. Panel A reports univariate statistics; Panel B reports the correlation coefficients among return volatility measured using the TSRV method of Zhang, Mykland, and Aït-Sahalia (2005); and Panel C reports the correlation coefficients among return volatility measured using the realized volatility approach. The sample consists of all 41 stocks that were included in the Dow Jones Industrial Average at some point from 01/01/2003 to 12/31/2014. The volatility measures for 15-minute returns reported in Panels A and B are obtained using the TSRV method applied to stock returns at 30, 45, and 60 second intervals, respectively, offset at 1 second intervals. The volatility measures for 15-minute returns reported in Panel C are obtained using the realized volatility method applied to stock returns at 30, 45, and 60 second intervals.

Panel A: Univariate statistics

	Mean	Median	Std. dev.
Stock return	5.63e-06	0.00	0.0030
Return volatility			
30 seconds	0.0025	0.0019	0.0021
45 seconds	0.0024	0.0019	0.0021
60 seconds	0.0023	0.0018	0.0020
Number of observations:	2796632		

Panel B: Correlation of volatility measures obtained using TSRV method

	Time interval for TSRV measure		
	30 seconds	45 seconds	60 seconds
30 seconds	1.000	0.994	0.984
45 seconds	0.994	1.000	0.997
60 seconds	0.984	0.997	1.000

Panel C: Correlation of realized volatility measures

	Time interval for returns		
	30 seconds	45 seconds	60 seconds
30 seconds	1.000	0.934	0.947
45 seconds	0.934	1.000	0.932
60 seconds	0.947	0.932	1.000

Table 2: Power law coefficients from data and nonparametric model

Calculations based on Consolidated Trades data from the TAQ database. The table compares the power law coefficients obtained from the returns data with those obtained from the conditional normal model with a nonparametric volatility distribution for each of the 41 stocks that were included in the Dow Jones Industrial Average at some point from 01/01/2003 to 12/31/2014. The second column reports the power law coefficient for stock returns obtained from the 15-minute return data, the third column reports the corresponding 95 percent confidence interval, and the fourth to sixth columns report the coefficients obtained from the conditional normal model with volatility within 15 minutes measured using the TSRV method of Zhang, Mykland, and Ait-Sahalia (2005). The latter results are obtained by applying the TSRV method to stock returns at 30, 45, and 60 second intervals offset at 1 second intervals, respectively. \* and \*\*, respectively, denote observations for which the model implied power law coefficient falls outside the 95 and 99 percent confidence intervals for the power law coefficient from the return data.

Stock	Estimated power law coefficient	Confidence interval	Power law coefficient from model		
			30 seconds	45 seconds	60 seconds
AA	2.93	[2.79, 3.06]	2.95	2.93	2.90
AIG	2.29	[2.14, 2.45]	2.37	2.33	2.35
AXP	2.60	[2.48, 2.72]	2.62	2.59	2.60
BA	3.17	[3.03, 3.32]	3.06	3.09	3.09
BAC	2.28	[2.17, 2.38]	2.35	2.32	2.31
C	2.21	[2.11, 2.31]	2.36**	2.31	2.27
CAT	3.01	[2.87, 3.14]	2.92	2.94	2.95
CSCO	3.26	[3.11, 3.41]	3.28	3.28	3.26
CVX	3.00	[2.87, 3.14]	2.92	2.92	2.92
DD	3.10	[2.96, 3.24]	3.01	3.03	3.02
DIS	3.08	[2.93, 3.22]	3.02	3.03	3.03
EK	3.05	[2.88, 3.23]	3.26*	3.21	3.18
GE	2.50	[2.38, 2.61]	2.56	2.53	2.53
GM	2.09	[1.96, 2.22]	2.34**	2.30**	2.25*
GS	2.50	[2.39, 2.62]	2.51	2.50	2.50
HD	3.07	[2.92, 3.21]	3.00	3.00	2.98
HON	3.10	[2.96, 3.25]	3.01	3.02	3.02
HPQ	3.26	[3.11, 3.41]	3.22	3.22	3.22
IBM	2.93	[2.79, 3.06]	2.82	2.84	2.83
INTC	3.46	[3.30, 3.62]	3.38	3.38	3.38

Contd.



Table 2(contd.): Power law coefficients from data and nonparametric model

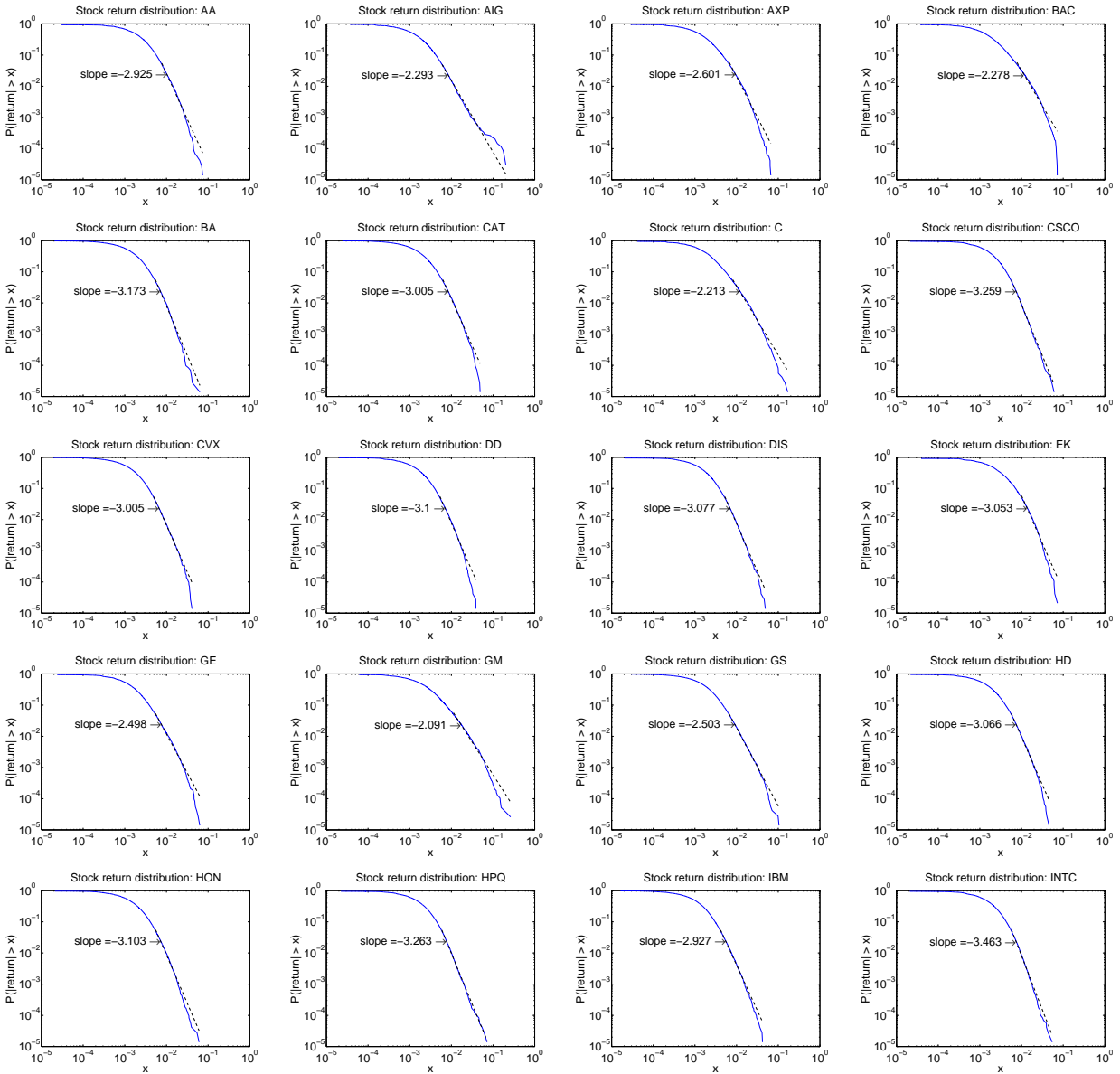
Stock	Estimated power law coefficient	Confidence interval	Power law coefficient from model		
			30 seconds	45 seconds	60 seconds
IP	2.92	[2.79, 3.05]	2.83	2.84	2.83
JNJ	3.00	[2.86, 3.14]	3.04	3.03	3.02
JPM	2.59	[2.47, 2.71]	2.59	2.58	2.56
KFT	3.10	[2.94, 3.26]	3.08	3.11	3.12
KO	3.17	[3.02, 3.32]	3.11	3.13	3.12
MCD	3.18	[3.03, 3.33]	3.04	3.06	3.08
MMM	3.06	[2.92, 3.21]	3.00	3.01	3.01
MO	2.76	[2.64, 2.89]	2.95**	2.95**	2.93**
MRK	2.93	[2.79, 3.06]	2.94	2.94	2.94
MSFT	3.16	[3.01, 3.31]	3.26	3.23	3.19
NKE	3.14	[3.00, 3.29]	3.05	3.07	3.10
PFE	3.20	[3.05, 3.35]	3.29	3.27	3.25
PG	3.08	[2.94, 3.22]	3.02	3.03	3.04
T	2.89	[2.75, 3.02]	2.92	2.91	2.90
TRV	2.63	[2.48, 2.78]	2.57	2.57	2.57
UNH	2.78	[2.65, 2.91]	2.76	2.79	2.81
UTX	3.15	[3.00, 3.29]	2.98*	2.98*	2.98*
V	2.83	[2.65, 3.00]	2.86	2.87	2.87
VZ	2.97	[2.84, 3.11]	2.96	2.96	2.94
WMT	3.21	[3.06, 3.35]	3.09	3.09	3.09
XOM	3.00	[2.86, 3.14]	2.96	2.96	2.95

Table 3: Summary statistics for normalized returns

Calculations based on Consolidated Trades data from the TAQ database. The table reports summary statistics for the normalized returns over 15-minute intervals for all the stocks in the sample, pooled together. Normalized returns are constructed by scaling stock returns with the corresponding volatility over that interval. The sample comprises the 41 stocks that were included in the Dow Jones Industrial Average at some point from 01/01/2003 to 12/31/2014. The volatility measures for 15-minute returns reported in Panels A and B are obtained using the TSRV method applied to stock returns at 30, 45, and 60 second intervals, respectively, offset at 1 second intervals.

	Time interval for TSRV measure		
	30 seconds	45 seconds	60 seconds
Mean	0.002	0.002	0.002
Standard deviation	0.957	0.940	0.916
Skewness	0.004	0.005	0.007
Kurtosis	2.786	2.898	3.055

Figure 1: Tail distribution of stock returns



Calculations based on Consolidated Trades data from the TAQ database. The figure plots the tail distribution of the absolute value of 15-minute stock returns from 01/01/2003 to 12/31/2014 for each of the stocks in the sample. The x-axis plots the absolute return and the y-axis plots the probability that the absolute return exceeds the given value. Both axes are on log scales. The figures also report the slope and fitted line obtained from a regression of the log tail probability on the log absolute return estimated on returns exceeding the 95th percentile of the empirical distribution for the absolute value of stock returns.

Figure 1(contd.): Tail distribution of stock returns

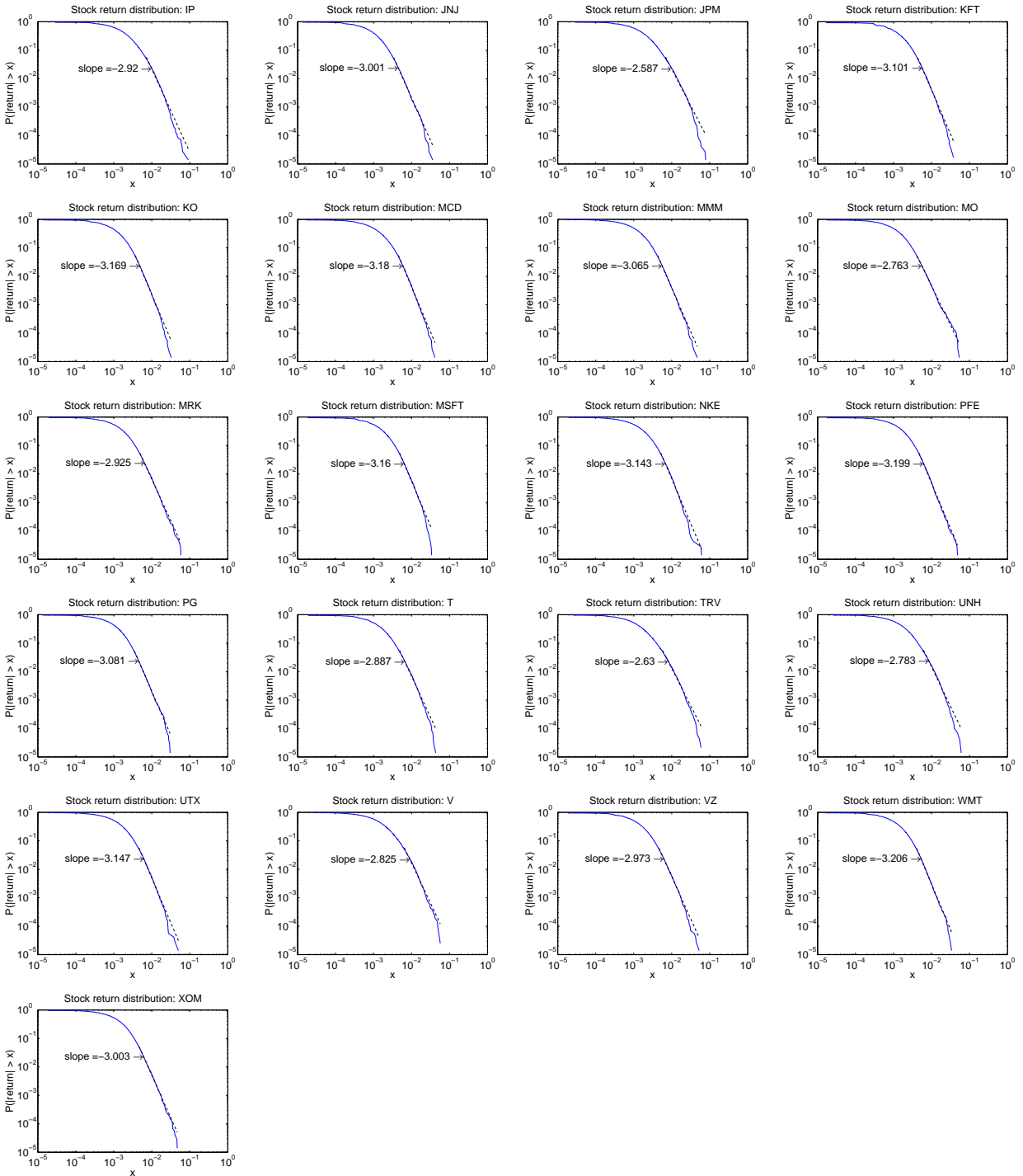
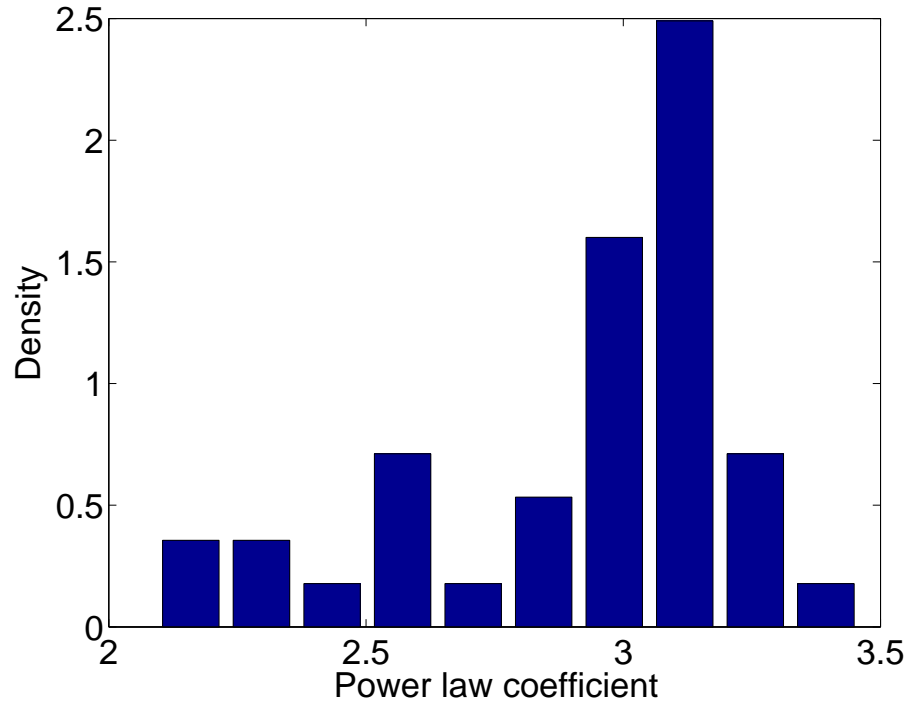


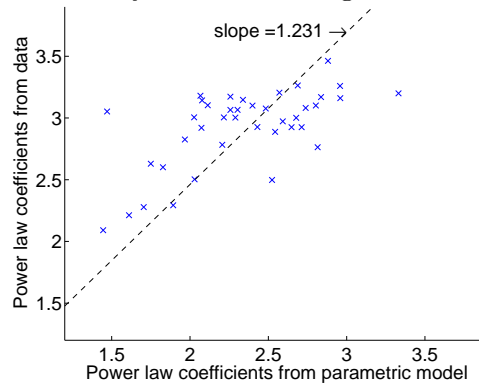
Figure 2: Distribution of power law coefficients



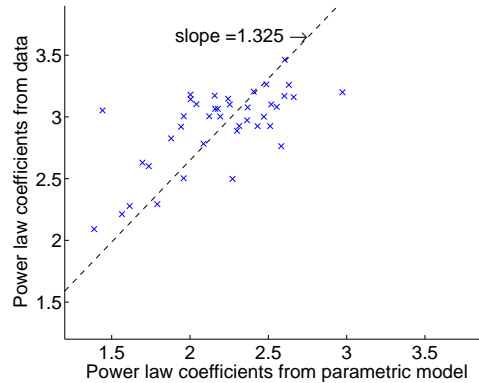
Calculations based on Consolidated Trades data from the TAQ database. The figure reports the cross-sectional distribution of the power law coefficients for stock returns at 15 minute intervals for the 41 stocks that were, at some point, included in the Dow Jones Industrial Average from 01/01/2003 to 12/31/2014.

Figure 3: Comparison of power law coefficients: Data and parametric model

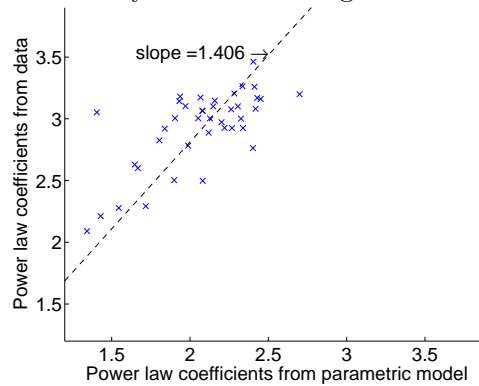
Panel A: Volatility measured using 30-second returns



Panel B: Volatility measured using 45-second returns

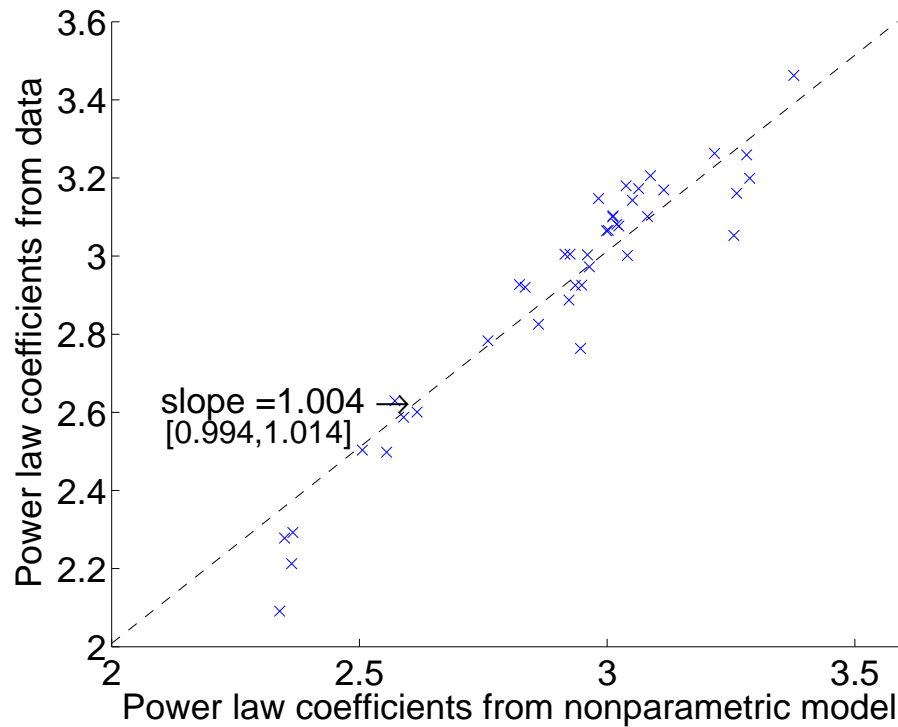


Panel C: Volatility measured using 60-second returns



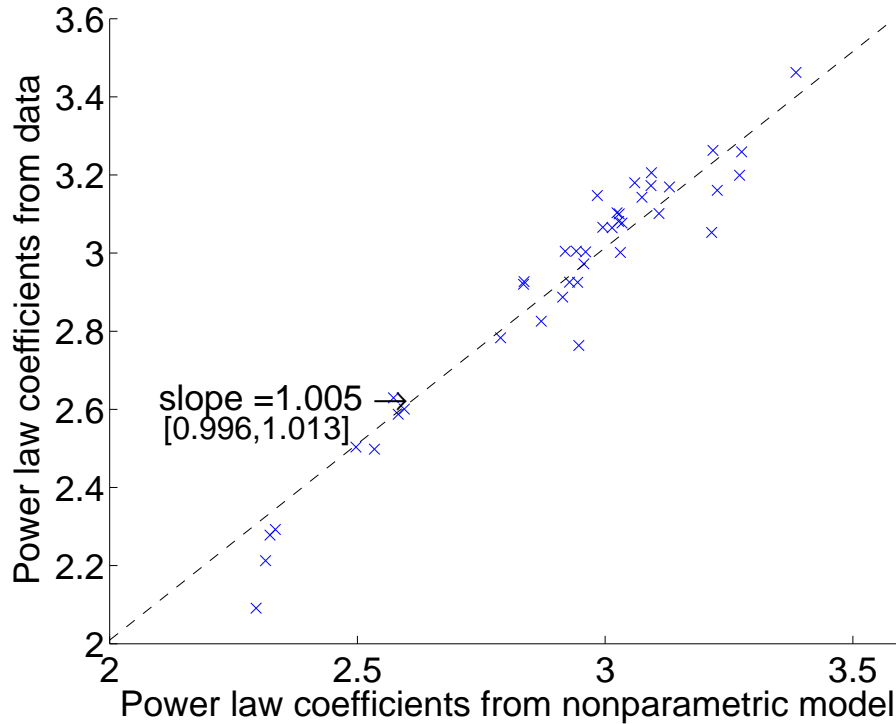
Calculations based on Consolidated Trades data from the TAQ database. The y-axes of the scatter plots present the power law coefficient for stock returns over 15 minute intervals; the x-axes present twice the shape coefficient of the inverse-gamma distribution for stock return volatility within 15 minute intervals. The sample comprises the 41 stocks that were, at some point, included in the Dow Jones Industrial Average from 01/01/2003 to 12/31/2014. Panels A, B, and C, respectively, measure volatility using the TSRV method of Zhang, Mykland, and Ait-Sahalia (2005) applied to returns over 30, 45, and 60 seconds offset at 1 second. The dashed lines represent OLS estimates of a regression through the origin.

Figure 4A: Comparison of power law coefficients: Data and nonparametric model



Calculations based on Consolidated Trades data from the TAQ database. The y-axis of the scatter plot present the power law coefficient for stock returns over 15 minute intervals; the x-axis present the power law coefficient obtained from simulating the conditional normal model with a nonparametric volatility distribution for stock return volatility within 15 minute intervals. The sample comprises the 41 stocks that were, at some point, included in the Dow Jones Industrial Average from 01/01/2003 to 12/31/2014. Volatility is measured using the TSRV method of Zhang, Mykland, and Ait-Sahalia (2005) applied to returns over 30 second intervals offset at 1 second. The dashed line represents the OLS fit of a regression through the origin of the power law coefficient from the data on the power law coefficient from the simulated nonparametric model. The square brackets report the 95 percent confidence interval for the slope coefficient of the regression.

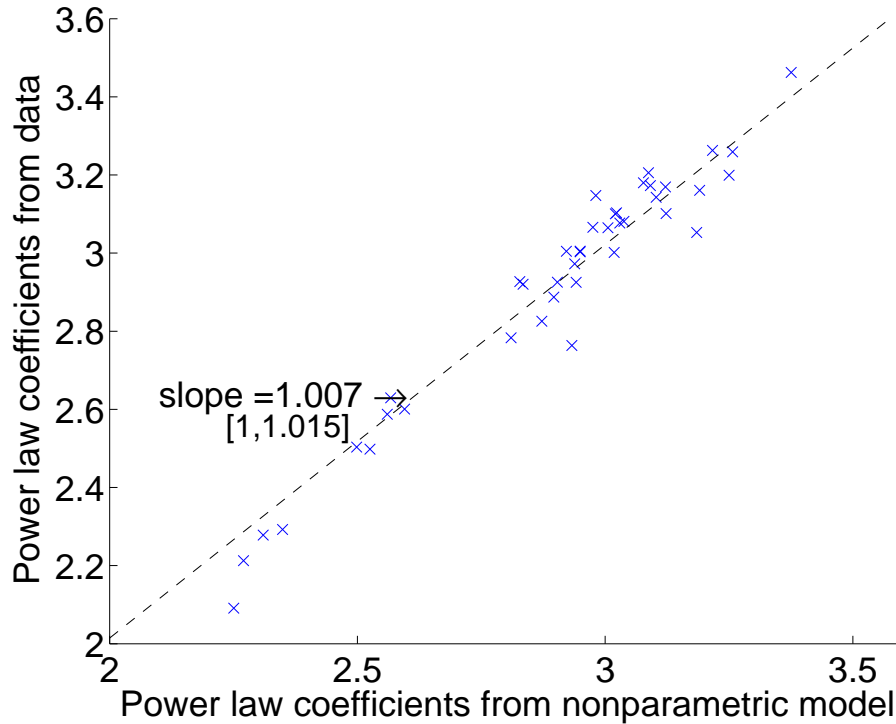
Figure 4B: Comparison of power law coefficients: Data and nonparametric model



Calculations based on Consolidated Trades data from the TAQ database. The y-axis of the scatter plot presents the power law coefficient for stock returns over 15 minute intervals; the x-axis presents the power law coefficient obtained from simulating the conditional normal model with a nonparametric volatility distribution for stock return volatility within 15 minute intervals. The sample comprises the 41 stocks that were, at some point, included in the Dow Jones Industrial Average from 01/01/2003 to 12/31/2014. Volatility is measured using the TSRV method of Zhang, Mykland, and Ait-Sahalia (2005) applied to returns over 45 second intervals offset at 1 second. The dashed line represents the OLS fit of a regression through the origin of the power law coefficient from the data on the power law coefficient from the simulated nonparametric model. The square brackets report the 95 percent confidence interval for the slope coefficient of the regression.



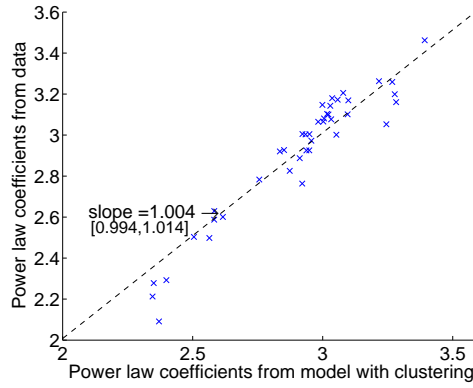
Figure 4C: Comparison of power law coefficients: Data and nonparametric model



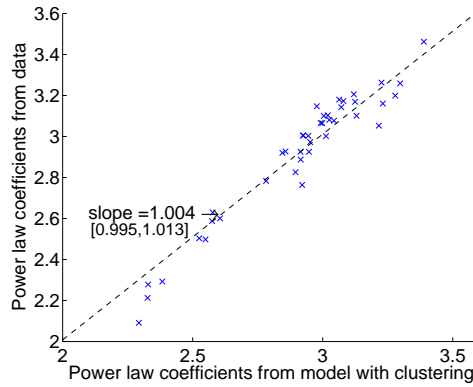
Calculations based on Consolidated Trades data from the TAQ database. The y-axis of the scatter plot presents the power law coefficient for stock returns over 15 minute intervals; the x-axis presents the power law coefficient obtained from simulating the conditional normal model with a nonparametric volatility distribution for stock return volatility within 15 minute intervals. The sample comprises the 41 stocks that were, at some point, included in the Dow Jones Industrial Average from 01/01/2003 to 12/31/2014. Volatility is measured using the TSRV method of Zhang, Mykland, and Ait-Sahalia (2005) applied to returns over 60 second intervals offset at 1 second. The dashed line represents the OLS fit of a regression through the origin of the power law coefficient from the data on the power law coefficient from the simulated nonparametric model. The square brackets report the 95 percent confidence interval for the slope coefficient of the regression.

Figure 5: Comparison of data and nonparametric model with volatility clustering

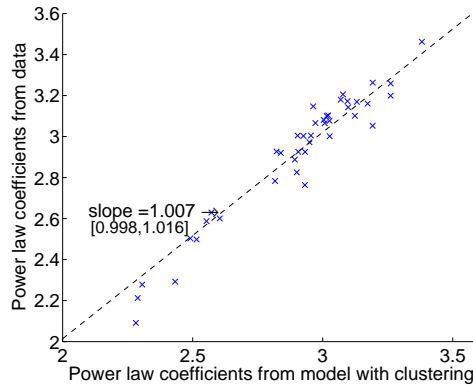
Panel A: Volatility measured using 30-second returns



Panel B: Volatility measured using 45-second returns



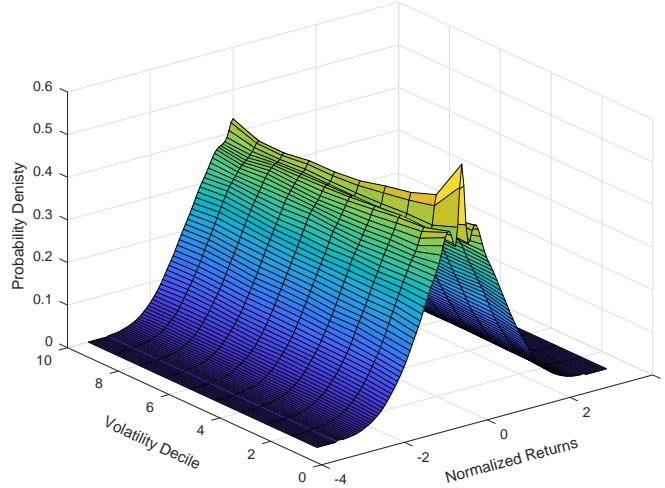
Panel C: Volatility measured using 60-second returns



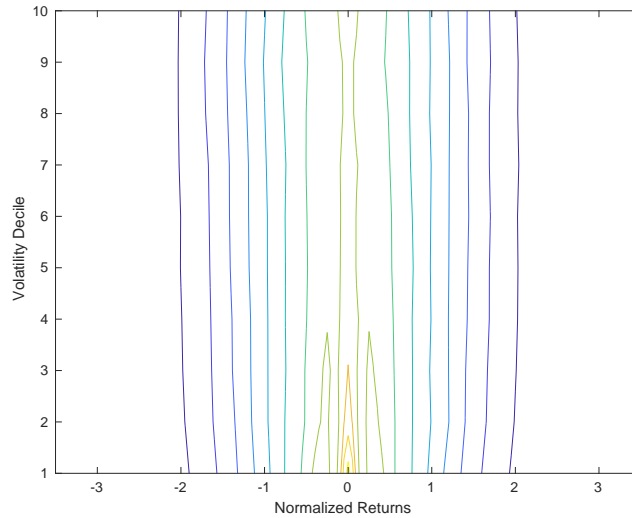
Calculations based on Consolidated Trades data from the TAQ database. The y-axes of the scatter plots present the power law coefficient for 15-minute stock returns; the x-axes present the power law coefficient obtained from simulating a nonparametric conditional normal model with clustering. The sample comprises the 41 stocks that were, at some point, included in the Dow Jones Industrial Average from 01/01/2003 to 12/31/2014. Panels A, B, and C, respectively, measure volatility using the TSRV method of Zhang, Mykland, and Ait-Sahalia (2005) applied to returns over 30, 45, and 60 seconds offset at 1 second. The dashed line represents the OLS fit of a regression through the origin.

Figure 6: Joint distribution of volatility and normalized returns

Panel A: Joint density of volatility and normalized returns



Panel B: Volatility and normalized return contours



Calculations based on Consolidated Trades data from the TAQ database. The figure presents the joint empirical distribution of volatility and normalized returns over 15 minute intervals. The volatility data are sorted into ten deciles. Panel A presents the density function for normalized returns for the volatility deciles. Panel B presents the corresponding contour lines for the density of normalized returns over volatility deciles. The sample comprises the 41 stocks that were, at some point, included in the Dow Jones Industrial Average from 01/01/2003 to 12/31/2014. Volatility with a 15-minute interval is measured using the TSRV method of Zhang, Mykland, and Ait-Sahalia (2005) applied to returns over 60 seconds offset at 1 second.