Journal of Empirical Finance 49 (2018) 123–141

Contents lists available at ScienceDirect

Journal of Empirical Finance

journal homepage: www.elsevier.com/locate/jempfin

Time-varying volatility and the power law distribution of stock returns

Missaka Warusawitharana

Board of Governors of the Federal Reserve System, Division of Supervision and Regulation, Mail Stop 1814, 20th and C Street NW, Washington, D.C., 20551, United States

ARTICLE INFO

JEL classification:
C58
D30
G12

Keywords:
Tail distributions
High frequency returns
Power laws
Time-varying volatility

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.[1] 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 et al. (1999), among others.

Gabaix et al. (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

E-mail address: missaka.n.warusawitharana@frb.gov.

[1] Gopikrishnan et al. (2000) find that tail distribution of returns on stock indices exhibit a power law with a power coefficient around 3; Plerou et al. (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 et al. (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.

https://doi.org/10.1016/j.jempfin.2018.09.004
Received 16 December 2016; Received in revised form 5 August 2018; Accepted 18 September 2018
Available online 1 October 2018
0927-5398/Published by Elsevier B.V.

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 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 s 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 min 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 et al. (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 et al. (2003) and Barndorff-Nielsen and Shephard (2002) runs into difficulty with market microstructure noise when applied at such high frequencies.[2] As such, the two-scales realized volatility method developed by Zhang et al. (2005) is used to obtain return volatility within 15 min intervals, the primary time interval used in the study. Applying this method to returns at 30, 45, or 60 s intervals offset at intervals of 1 s each, one obtains robust measures of return volatility at 15 min 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 distribution of returns 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 min 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.[3]

[2] Barndorff-Nielsen and Shephard (2004) extends this method to estimate realized covariances and regression coefficients.

[3] Without loss of generality, we will assume that high-frequency stock returns have zero mean.