

Heavy Tails and Mixed Distribution Hypothesis

by

Jonathan B. Hill*

Dept. of Economics

University of North Carolina - Chapel Hill

Abstract

We outline the Mixed Distribution Hypothesis as a means to explain heavy tails in financial time series. We discuss the hypothesis' historical roots, and fully present the most popular, and original, form of the hypothesis and its implications for modeling asset returns. Original contributions and modern extensions are cited.

1. INTRODUCTION The mixed distribution hypothesis [MDH] in finance maintains heavy tails in price fluctuations $\Delta X(t)$ can be explained by embedding

*Address: Dept. of Economics CB 3205, University of North Carolina-Chapel Hill, Chapel Hill, NC 27599-3305; Email : jbhill@email.unc.edu; www.unc.edu/~jbhill.

Key words : mixtures of distributions; heavy tails; price-volume relationship.

randomized time $\xi(t)$. Intuitively, the net price change $\Delta X(t)$ over period t is hypothesized to be the sum of $\xi(t)$ incremental inter-period steps, say $\Delta X_{t,i}$, giving

$$\Delta X(\xi(t)) = \sum_{i=1}^{\xi(t)} \Delta X_{t,i}.$$

Randomized time $\xi(t)$ can be thought of as the amount of economic time or the number of information events that traders actually experience during t , typically associated with volume traded $V(t)$. The resulting convolution $\Delta X(\xi(t))$ has a mixed distribution that can capture leptokurtosis, explain heterogeneity, including conditional heteroscedasticity, and improve model fitting similar to nonparametrics. Applications abound in finance (equity returns, bond prices) and macroeconomics (income, unemployment, exchange rates). See Section 2.

The statistical foundations of mixtures are well known (e.g. [7]). If $y(t) := \Delta X(t)$ has density $f_y(y; \xi)$ and economic time $\xi(t) \geq 0$ has density $f_\xi(\xi)$, then the observed convolution $y(\xi(t))$ has a density mixture

$$f_{y(\xi)}(y) = \int_0^\infty f_y(y; \xi) f_\xi(\xi) d\xi. \quad (1)$$

See [22] for a survey of mixture models, including generalized gamma and beta, log-Cauchy and log-normal.

The use of compound or mixed distributions to explain leptokurtosis in price fluctuations $\Delta X(t)$, while retaining a finite variance, first appears in the finance literature in the 1960's¹. A sequence of papers, including [26], [25], and prominently Clark [4],

¹Evidently Newcomb [24] first used a mixture of normal distributions to account for numerous

exploited in various way a subordinated finite variance process model $\Delta X(\xi(t))$ as a direct response to Mandelbrot's [20] controversial α -stable hypothesis of aggregate prices. The latter "extreme value theory" framework permits infinite variance, but at a cost of estimation and inference challenges (see eqf02_009 [5] and eqf19_017 [17] in this volume). Nevertheless, only the α -stable laws form a closed distribution class under affine transformations, finite variance α -stable laws are the Gaussian laws which are necessarily symmetric, and infinite variance α -stable laws permit asymmetry, making this class attractive for applied statistical analysis in finance (see [28] for an encyclopedic treatment of the α -stable laws).

Tests of MDH against non-mixtures, including normality or α -stability, have mixed results ([21], [30], [12]), and there exists a sharp identifiability problem that is typically ignored (consider [19]). The subordinated process premise has successfully evolved into randomized volatility ([1], [3], [11]) which now receives vigorous attention in the finance literature ([8], [10], [11], [27]) and extreme value theory literature ([14], [15], [18]), and some applications of distribution mixtures in finance are motivated by specification flexibility rather than explaining heavy tails (e.g. [9]).

In the remainder of this note we discuss at length Clark's [4] seminal model, we note the early attempts to improve on it, and conclude with remarks on recent extensions in the stochastic volatility literature.

"discordant" observations (i.e. heavy tails), with an application to planetary orbits.

2. RANDOMIZED TIME AND SUBORDINATION Clark's [4] influential model of asset returns exploits Bochner's [2] and Feller's [7] work on subordinated Gaussian processes. Denote by μ_y , σ_y^2 and κ_y respectively the mean, variance and kurtosis of y , and $\sigma_{y|z}^2(t)$ is the conditional variance given random z .

If $\{X(t)\} \equiv \{X(t) : -\infty < t < \infty\}$ is a stochastic process and $\{\xi(t)\}$, $\xi : \mathbb{N} \rightarrow \mathbb{R}_+$, is a "driving process" then the stochastically indexed $X(\xi(t))$ is "subordinated" to $X(t)$. If $\{\Delta\xi(t)\}$ is stationary and independent with mean $\mu_{\Delta\xi} > 0$, and $\Delta X(t) \stackrel{iid}{\sim} (0, \sigma_{\Delta X}^2)$, then the subordinated fluctuations $\Delta X(\xi(t))$ are stationary and independent, and distributed

$$\Delta X(\xi(t)) \stackrel{iid}{\sim} (0, \mu_{\Delta\xi} \times \sigma_{\Delta X}^2).$$

Since $\mu_{\Delta\xi}$ reflects the arrival of new information, in Clark's model only the average amount of new information effects the dispersion of imbedded returns $\sigma_{\Delta X(\xi)}^2$. If in addition price fluctuations are normally distributed $\Delta X(t) \stackrel{iid}{\sim} N(0, \sigma_{\Delta X}^2)$ and $\sigma_{\Delta\xi}^2 < \infty$ and $\Delta\xi(t)$ are independent of $X(t)$, then price fluctuations $\Delta X(\xi(t))$ have kurtosis

$$\kappa_{\Delta X(\xi)} = 3 \left(1 + \frac{\sigma_{\Delta\xi}^2}{\mu_{\Delta\xi}^2} \right) > 3$$

a simple positive function of $\sigma_{\Delta\xi}^2$, and larger than 3, the benchmark value for normals and therefore for what is considered canonically to be heavy-tailed. The tails of $y \sim (\mu, \sigma^2)$ implies y is distributed with mean μ and variance σ^2 . *iid* is short for independent and identically distributed.

price fluctuations $\Delta X(\xi(t))$ grow heavier as the variance $\sigma_{\Delta\xi}^2$ of information flow increases, and are heavy-tailed in the strict sense of being thicker than normals for any directing process $\{\xi(t)\}$ with $\sigma_{\Delta\xi}^2 > 0$. Further, as in [4], if iid $\Delta\xi(t)$ are lognormal with parameters $\{c, v^2\}$ then

$$\mu_{\Delta\xi} = \exp\{c + v^2/2\} \text{ and } \sigma_{\Delta\xi}^2 = \exp\{2c + v^2\} \times [\exp\{v^2\} + 1],$$

and from (1) the convolution $y(\xi(t)) \equiv \Delta X(\xi(t))$ is lognormal-normally distributed:

$$f_y(y) = \frac{1}{2\pi\sigma_{\Delta X}^2\sigma_{\Delta\xi}^2} \int_0^\infty u^{-3/2} \exp\left\{\frac{-(\ln u - c)^2}{2v^2} - \frac{y^2}{2u\sigma_{\Delta X}^2}\right\} du.$$

Volume traded $V(t)$ may be positively associated with economic $\xi(t)$, in which case equity returns should be heavier tailed during periods of substantial trader activity. Heimstra and Jones [13], for example, use this prediction to test for variance causality from volume to returns under a finite variance assumption.

Many models for relating randomized time to volume have been suggested. See [16] for a survey. Clark [4] uses the parametric specification

$$\sigma_{\Delta X|\xi}^2(t) = \beta V(t)^\gamma, \text{ where } \beta, \gamma > 0,$$

for the log-price $X(t)$ of cotton future prices. Epps and Epps [6] use a parsimonious parametric asset price model to support Clark's mixtures of distribution model. Tauchen and Pitts [29] propose an asset price model that relates price fluctuations to volume via a variance-components framework with trader-specific and trader-wide

shocks. Unlike Clark their model predicts $\Delta X(t)$ and $V(t)$ are independent for each transaction, but have moments with common parameters. For example, $\sigma_{\Delta X}^2$ and μ_V are both increasing functions of the variance of the trader-specific shock and therefore exhibit spurious positive association (as opposed to Clark's ad hoc association).

Stochastic volatility models for asset returns inherently exploit the premise that a mixture of distributions captures the arrival of news, making this literature the latest generation of Clark's model (see [1], [11], [27]). The Brownian semimartingale framework for stochastic volatility has become dominant in the option pricing literature by permitting volatility clustering and accounting for market incompleteness (which abnegates option redundancy), providing flexible extensions of the Black-Scholes option pricing model. See especially eqf19_006 [8] and eqf19_001 [27] in this volume, and consult [10] and [23].

REFERENCES

- [1] Anderson T. (1996). Return Volatility and Trading Volume: An Information Flow Interpretation of Stochastic Volatility, *Journal of Finance* **51**, 169-204.
- [2] Bochner, S. (1960). *Harmonic Analysis and the Theory of Probability*; University of California Press, Berkeley.
- [3] Carr, P., H. Geman, D.B. Madan, M. Yor (2003). Stochastic Volatility for Lévy Processes, *Mathematical Finance* **13**, 345–382.
- [4] Clark, P.K. (1973) Subordinated Stochastic Process Model with Finite Variance

for Speculative Processes, *Econometrica* **41**, 133-153.

[5] Davis, R.A. (2009). Heavy Tailed Processes [eqf02_009], in *Encyclopedia of Quantitative Finance*, R. Cont (ed.); Wiley, New York.

[6] Epps, T.W. and M.L. Epps (1976). The Stochastic Dependence of Security Price Changes and Transaction Volumes: Implications for the Mixture-of-Distributions Hypothesis, *Econometrica* **44**, 305-321.

[7] Feller, W. (1971). *An Introduction to Probability Theory and its Applications, Vol. II*; Wiley, New York.

[8] Fouque, J.-P. (2009). Option Pricing with Stochastic Volatility [eqf19_006], in *Encyclopedia of Quantitative Finance*, R. Cont (ed.); Wiley, New York.

[9] Geweke, J. and M. Keane (2007). Smoothly Mixing Regressions, *Journal of Econometrics* **138**, 252-290.

[10] Garcia, R., E. Ghysels and E. Renault (2009). Econometrics of Option Pricing Models, *Handbook of Financial Econometrics*, Y. Aït-Sahalia and L.P. Hansen (eds.), North Holland, forthcoming.

[11] Ghysels, E., A. Harvey and E. Renault (1995). Stochastic Volatility, in: *Handbook of Statistics 15, Statistical Methods in Finance*, G.S. Maddala and C.R. Rao (eds.); North Holland, Amsterdam.

[12] Hall, J.A., W. Broersen, S.H. Irwin (1989). The Distribution of Futures Prices: A Test of the Stable Paretian and Mixture of Normals Hypotheses, *Journal of Financial*

and *Quantitative Analysis* **24**, 105-116.

[13] Heimstra, C., and J. D. Jones (1994). Testing for Linear and Nonlinear Granger Causality in the Stock Price-Volume Relation, *Journal of Finance* **49**, 1639-1664.

[14] Hill, J.B. (2008a). Nonparametric Tail Dependence: Probability Theory with an Application to Stochastic Volatility, Dept. of Economics, University of North Carolina - Chapel Hill, submitted.

[15] Hill, J.B. (2009). On Functional Central Limit Theorems for Dependent, Heterogeneous Arrays with Applications to Tail Index and Tail Dependence Estimation, *Journal of Statistical Planning and Inference*, forthcoming.

[16] Karpoff, J.M. (1987). The Relation between Price Changes and Trading Volume: A Survey, *Journal of Financial and Quantitative Analysis*, **22**, 109-126.

[17] Klüppelberg, C. (2009). Extreme Value Theory [eqf19_017], in *Encyclopedia of Quantitative Finance*, R. Cont (ed.); Wiley, New York.

[18] Klüppelberg, C., and A. Lindner (2008). Extremes of Random Volatility Models, Dept. of Mathematics, Universität München, submitted.

[19] Li, L.A., and N. Sedransk (1988). Mixtures of Distributions: A Topological Approach, *Annals of Statistics* **16**, 1623-1634.

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

[21] Mandelbrot, B. (1973). Comments on: "A Subordinated Stochastic Process

- Model with Finite Variance for Speculative Processes" by P.K. Clark, *Econometrica* **41**, 157-159.
- [22] McDonald, J.B. and R.J. Butler (1987). Generalized Mixture Distributions with an Application to Unemployment Duration, *Review of Economics and Statistics* **69**, 232-240.
- [23] Mykland, P. and L. Zhang (2008). Inference for Continuous Semimartingales Observed at High Frequency: A General Approach, mimeo, University of Chicago.
- [24] Newcomb, S. (1886). A Generalized Theory of the Combination of Observations so as to Obtain the Best Result, *American Journal of Mathematics* **8**, 343-366.
- [25] Praetz, P.D. (1972) The Distribution of Share Price Changes, *Journal of Business* **45**, 49-55.
- [26] Press, J.S. (1967). A Compound Events Model for Security Prices, *Journal of Business* **40**, 317-335.
- [27] Renault, E. (2009). Econometrics of Option Pricing [eqf19_001], in *Encyclopedia of Quantitative Finance*, R. Cont (ed.); Wiley, New York.
- [28] Samorodnisky, G. and M. Taqqu (1994). *Stable Non-Gaussian Random Processes*; Chapman and Hall, New York.
- [29] Tauchen, G.E., and M. Pitts (1983). The Price Variability-Volume Relationship on Speculative Markets, *Econometrica* **51**, 485-505.
- [30] Upton, D.E., and D.S. Shannon (1979). The Stable Paretian Distribution, Sub-

ordinated Stochastic Processes, and Asymptotic Lognormality: An Empirical Investigation, *Journal of Finance* **34**, 1031-1039.