Network Effects and Kurtosis

Ben Klemens

January 11, 2009

Abstract

It is well known that the distribution of returns from various financial instruments are leptokurtic, meaning that the distributions have ‘fatter tails’ than a Normal distribution, and have positive skew, meaning that one sees exceptional positive returns more often than exceptional negative returns. This paper presents a graceful micro-level explanation for these effects, using agents whose private valuations have Normally-distributed errors, but whose utility function includes a term for the percentage of others who are also buying a security.

Working Paper. The usual caveats apply. Please send comments to ben@klemens.org.

1 Introduction

Many researchers have pointed out that day-to-day returns on equities have ‘fat tails,’ in the sense that extreme events happen much more frequently than would be predicted by a Normal distribution, and have positive skew, meaning that extreme positive returns are more likely than extreme negative returns. This has been re-verified by many of the sources listed below.

Here, I present an explanation for the non-Normality of equity returns using a micro-level model where agents observe and emulate the behavior of others. There are several reasons for rational agents to take note of the actions of other rational agents; the model here is agnostic as to which best describe real-world agents, but given some motivation to emulate others, this is sufficient to explain the wider-than-Normal distribution of equity returns.

2 Literature

This section gives an overview of two threads of the economics literature that do not quite meet. The first is an overview of the existing literature on the distribution of equity returns; the second is a survey of the situations posited in the literature where individuals gain utility from emulating others.
2.1 Fitting non-Normal distributions

The Normal distribution has two parameters: the mean \( \mu \) and standard deviation \( \sigma \) (or equivalently, the square of standard deviation, the variance). The variance is the second central moment:

\[
\sigma^2 = \int_{-\infty}^{\infty} (x - \bar{x})^2 f(x) dx,
\]

where \( x \in \mathbb{R} \) is a random variable, \( \bar{x} \) is the mean of \( x \), and \( f(x) \) is the probability distribution on \( x \).

One could similarly define the third moment, known as the skew, and the fourth moment, the kurtosis:

\[
S = \int_{-\infty}^{\infty} (x - \bar{x})^3 f(x) dx, \quad \text{and}
\]

\[
\kappa = \int_{-\infty}^{\infty} (x - \bar{x})^4 f(x) dx.
\]

For a Normal distribution, \( \kappa = 3 \sigma^4 \). Let the normalized kurtosis be \( \kappa/\sigma^4 \); then the normalized kurtosis for a Normal distribution is three.\(^1\) One can use this fact to check empirical distributions for deviations from the Normal.

A Normal distribution is symmetric and therefore has zero skew, and this fact can also be used to check empirical distributions for deviations from the Normal distribution. Similarly to the kurtosis, the skew is customarily normalized by dividing by \( \sigma^3 \).

Fama [1965] ran such a test on equity returns, and found that they were leptokurtic, meaning that \( \kappa \gg 3 \sigma^4 \), and had positive skew. However, he is not the first to notice these features—Mandelbrot [1963, footnote 3] traces awareness of the non-Normality of return distributions as far back as 1915, and many of the papers cited below reproduce the results using their own data sets.

Most of the explanations for the deviation from the Normal have focused on finding a closed-form probability density function (PDF) that better fits the data. Mandelbrot [1963] showed that a stable Pareto (aka symmetric-stable) distribution fit better than the Normal. Blattberg and Gonedes [1974] showed that a renormalized Student’s \( t \) distribution fit better than a symmetric-stable distribution. Kon [1984] found that a sum of Normal distributions fit better than a Student’s \( t \). The mixture model produces an output distribution by summing a first Normal distribution \( \mathcal{N}(\mu_1, \sigma_1) \) with a separate second Normal distribution \( \mathcal{N}(\mu_2, \sigma_2) \). Depending on the values of the five input parameters (two means, two variances, and a mixing parameter), the distribution produced by summing the two can take on a wide range of mean, standard deviation, skew, and kurtosis.

\(^1\)Note that the definition of kurtosis is not quite standard: some define what I call the normalized kurtosis as simply the kurtosis, others subtract three, &c.
However, the mixture model raises a few critiques. First, Kon found that the sum of two distributions satisfactorily matches only about half of the equity return distributions he tests. Others require as many as four input distributions—and thus eleven input parameters—to explain the four moments of the distribution to be matched.

Also, as with all of the distribution models, the use of a sum of several distributions only raises the question of why the given distributions go beyond being a good fit and are a valid explanations of market behavior. After all, a Fourier transform can fit any data set to arbitrary precision, but it is not necessarily an explanation of market behavior. This brings us to the second thread of the literature, covering the micro-level behavior of market actors.

2.2 Emulation

The literature provides many rational motivations for emulating others, variously termed herding, information cascades, network effects, spillovers—not to mention simple questions of fashion. Unfortunately, none of these models were written with the intention of describing an observed leptokurtic distribution.

The restaurant problem Among the most common of the models where agents emulate others are the ‘herding’ or ‘information cascade’ models, e.g. Banerjee [1992] or Bikhchandani et al. [1992]. In this model, agents use the prior choices of other agents as information when making decisions. So if ten people are at restaurant A and none at restaurant B, an agent may choose A despite having private information advising for B.

This story can easily be shown to produce a bifurcated distribution of results: over many days, restaurant A will show either about 0% attendance or about 100% attendance every day. Many goods show such a blockbuster/flop dichotomy, such as movies [de Vane y and Walls, 1996].

But for our purposes, a sharply bimodal distribution is not desirable. First, one would be hard-pressed to find an equity whose returns are truly bimodal. Second, such a bifurcated outcome distribution is typically platykurtic, the opposite of the leptokurtosis we seek. Consider an ideal distribution with density $r$ at $a$ and density $1 - r$ at $b$ (for any values of $a, b \in \mathbb{R}$). The distribution has normalized kurtosis equal to

$$\frac{1}{r - r^2} - 3.$$ 

For a symmetric distribution, $r = 0.5$, the normalized kurtosis is one, and it remains less than three for any $r \in (.211, .789)$. Thus, a model that predicts a bifurcated distribution can only explain the data if the distribution is severely lopsided, which is not sustainable for equity returns—such a distribution would quickly send prices to zero or infinity.

Network externalities This is a property of goods where consumption by others increases the utility of the good, such as computer equipment that needs
to interoperate with others’ equipment, or coordination problems like the choice of whether to drive on the right or left side of the road. The typical analysis (e.g., that of Choi [1997]) is similar to that of the restaurant problem, except expected utility is gained directly instead of through information; the discussion above applies.

**Finance** There are a number of reasons for why the manager of an asset portfolio would emulate others:

- Pricing is partly based on the value of the underlying asset and partly on what others are willing to pay for the asset. At the extreme, people will buy a stock which pays zero dividends only if they are confident that there are other people who will also buy the stock; as more people are willing to buy, the value of the stock to any individual rises.

- It has long been a lament of the fund manager that if the herd does badly but he does well, he sees little benefit; but if the herd does well and he does badly, then he gets fired. Therefore, ‘behave like others’ may very explicitly appear in a fund manager’s utility function.

- Since an undercapitalized company is likely to fail, the success of a public offering may depend on how well-subscribed it is, providing another justification for putting the behavior of others in the fund manager’s utility function.

- If a large number of banks take simultaneous large losses, then they may be bailed out; since a bail-out is unlikely if only one bank takes a loss, this may also serve as an incentive for financiers to take risks together.

- Simply following the herd: “[...] elements such as fashion and sense of honour affected the banks’ decision to take part in a syndicated loan. Banks are certainly not insensitive to prevailing trends, and if it is ‘the in thing’ to take part in syndicated loans[...], people sometimes consent too readily.” [Jepma et al., 1996, p 337]

As noted above, the model here is a reduced form model which simply assumes that a financier’s expected utility from an action is increasing in the percentage of other people acting. I make no effort to explain which of the above motivations are present at any time, but posit that given these effects, the model below is applicable.

Empirical studies of analyst recommendations find that they do indeed herd. For example, Graham [1999] finds evidence of herding among investment newsletter recommendations, and finds that the more reputable ones are more likely to herd. Meanwhile, Hong et al. [2000] finds evidence of herding among investment analysts, and finds that inexperienced analysts are “more likely to be terminated for bold forecasts that deviate from consensus,” and therefore less reputable analysts are more likely to herd. Welch [2000] find that
an analyst recommendation has a strong impact on the next two recommendations for the same security by other analysts, and that this effect is uncorrelated with whether the recommendations prove to be correct or not. Although these papers disagree in the details, they all find empirical evidence that analysts are inclined to behave like other analysts (and therefore the people who listen to analysts are likely to also behave alike), so the model below is apropos.

Within the theoretical finance literature, papers abound regarding herding behavior (Grossman [1976, 1981], Radner [1979], Choi [1997], Minchert and Scotchmer [1999]), although they concern themselves not with explaining herding, but with the information aggregation issues entailed by herding.

### 3 The model

The model consists of a plurality of agents (the simulations below use 10,000), each privately deciding whether to purchase a good. Each has an individual taste for consuming, \( t \in \mathbb{R} \), where \( t \sim \mathcal{N}(\epsilon, 1) \); \( \epsilon \) is a small positive offset, fixed at zero and 0.05 in the simulations to follow.\(^2\) Let the desire to emulate others be represented by \( \alpha \in [0, \infty) \). Let the percentage of the population consuming be \( k \in [0, 1] \). Then the utility from consuming is simply

\[
U_c = t + \alpha k.
\]

The utility from not consuming is

\[
U_{nc} = \alpha(1 - k).
\]

That is, agents who do not consume get utility from emulating the \( 1 - k \) agents who also do not consume, but have a taste for non-consumption normalized to zero. One can show that this normalization is without loss of generality. Agents consume iff \( U_c > U_{nc} \).

A Bayesian Nash equilibrium is a set of acting agents, comprising \( k_a \) percent of the population, where all acting agents have \( U_c > U_{nc} \) given \( k_a \) percent acting, and all agents outside the acting set have \( U_c \leq U_{nc} \) given \( k_a \) percent acting.

It can be shown that, given the assumptions here, the game has a cutoff-type equilibrium, where there is a cutoff value \( T \) such that every agent with private tastes greater than \( T \) acts and every agent with \( t \leq T \) does not act. One can show that there is exactly one equilibrium to every game.\(^3\)

\(^2\)The value of \( \epsilon \) is basically irrelevant for the normalized kurtosis measure (as the number of runs goes to infinity), but its sign is relevant for skew: if \( \epsilon \) is positive, the result is positive skew as presented here, but if \( \epsilon \) is negative, the entire story reverses and there is increasingly negative skew with larger values of \( \alpha \). Because equity returns have historically been long-run positive, one can expect that the mean private valuation is positive for the equities studied in the literature, and skew will be positive.

\(^3\)Brock and Durlauf [2001] assume a cutoff distribution for a situation such as the one here, but the existence of a unique cutoff can be derived from the assumptions either in their paper or here.
3.1 Implementation

The algorithm for a single run of the simulation is displayed in Figure 1. In each step, agents consume or do not based on the value of $k$ from the last step, and the process repeats until the value of $k$ no longer changes. The algorithm simply solves for equilibrium via tatonnement—the equilibrium reached via market simulation is exactly the unique Nash equilibrium described above. The algorithm is computationally simpler than searching for the cutoff, and has the added benefit of being a direct representation of how a market goes about finding an equilibrium.

\begin{itemize}
  \item Fix $n$.
  \item Generate a new population of agents:
    \begin{itemize}
      \item Each has a taste $t$, drawn from a $N(\epsilon, 1)$ distribution.
      \item Each has an even chance of initially consuming or not.
    \end{itemize}
  \item Fix the initial proposed value of $k$ at \( \frac{1}{2} \).
  \item while $k$ this period is not equal to $k$ last period:
    \begin{itemize}
      \item for each agent:
        \begin{itemize}
          \item consume if $U_k > U_{nk}$.
        \end{itemize}
      \end{itemize}
  \item Recalculate $k$
\end{itemize}

Figure 1: The algorithm for finding the equilibrium level of consumption

The code itself was written in C using the Apophenia library, described in Klemens [2008], and is available upon request.

The equilibrium is a deterministic value of $\alpha$ and the 5,000 draws of individual preferences, but given that the preferences are drawn at random, different runs based on identical parameters will produce a different equilibrium cutoff $T$ and percent acting $k$.

3.2 Results

Begin with the symmetric case, where $\epsilon = 0$, so agents’ private tastes are drawn from a $N(0, 1)$ distribution. Figure 2 shows a sequence of distributions, from the distribution of $k$ given $\alpha = 0$ up to the distribution for $\alpha = 4$. Small values of $\alpha$ give a Normal output distribution of prices; large values give a coordination-game style bifurcation; and there is a small range where the transition occurs, and the distribution is neither bifurcated nor Normal.

The small transition range is especially clear when we look at the kurtosis of each $\alpha$'s distribution. As in Figure 3, the normalized kurtosis is three for small values of $\alpha$ (as for a Normal distribution), has a quick period of transition, and is then one for large values of $\alpha$ (as for a bimodal distribution).

However, the bifurcation is fragile: if $\epsilon > 0$, then the tail of the distribution where no one buys the good disappears. Figure 4 shows the sequence of distributions where $\epsilon = 0.05$. For $\alpha = 0$, the distribution matches the input
Figure 2: Above, a sequence of distributions, for $\alpha = 0$ in front up to $\alpha = 2$ at the back. Vertical axis is the number of runs (out of $\sim 20,000$ per $\alpha$) in the given histogram bin. Three slices are shown at bottom.
distribution of tastes, with a cutoff price near zero and a standard Normal distribution around that peak. At $\alpha = 4$, the density of the distribution is tightly packed around four standard deviations below zero (basically, the point where one would be indifferent with 100% acting).

But, just as there was not a steady transition from the Normal regime to the bifurcated regime for $\alpha = 0$, there is no such steady transition here. The normalized kurtosis again reveals this.

Figure 5 plots $\kappa/\sigma^4$ for each $\alpha$. The neighborhood of $\alpha \approx 1.3$ is again salient, because the normalized kurtosis in that range is significantly larger than three.

As shown in the bottom plot of Figure 5, normalized skew follows the same story relative to $\alpha$ as did kurtosis: it peaks around 1.3, although it never quite falls back to zero.

Discussion The story is not as simple as saying that an emulative utility function leads to an outcome distribution with high normalized kurtosis.

A herd of blind followers (where $\alpha$ is very large) produce a tight distribution of outcomes, with no fat tails. On the other end, the outcome distribution matches the mesokurtic input distribution. Only in the mid-range, when both forces are in balance, one gets the sort of outcomes observed in equity data: returns to some extent mirror the Normally-distributed set of beliefs, but there are occasional herding outcomes, wherein an extreme outcome occurs. In such a situation, where agents care about their private information but still keep an eye on where others are going, the overall distribution matches what we observe.
Figure 4: Two views of the $\alpha$-cutoff-frequency relation. The surface is a series of one PDF for each value of $\alpha$. Three of these PDFs are displayed in 2-D form at bottom.
Figure 5: The relationship between $n$ (on the horizontal axis) and $\kappa/\sigma^4$ (on the vertical axis, top) or $S/\sigma^3$ (on the vertical axis, bottom).
4 Conclusion

There are several explanations for why rational agents would choose to emulate others, all of which advise that a utility function meant to describe a trader in the finance markets should include a term for the desire to emulate others.

Meanwhile, we know that equity return distributions show a number of deviations from the Normal distribution implied by naïve application of a Central Limit Theorem. Adding a term for the emulation of others to individual utilities produces aggregate outcome distributions that show these same deviations from Normal: extreme outcomes happen more often, and do so asymmetrically. Thus, models of market volatility can better mirror real-world outcomes by including an emulative component in the model.

References


