Basing Industry Average
(Industry Benchmark) on a "thick weighted median"

We have developed our own method for computing Industry Averages, data-points previously supplied to us, pre-computed based on a proprietary algorithm, by our former data vendor.

At first glance, this might seem an easy task. We all know what a mathematical average is. We also know what a median is.

In the real world, the median is usable (and some members may create this on their own via the FMedian function), but not ideal. By zeroing in on the center of the distribution, numbers applicable to very small companies, even those that trade rarely as penny stocks, are every bit as influential as the mega-sized blue-chips.

The mathematical average, on the other hand, is completely unusable as is the frequent theoretician's favorite, the market-capitalization weighted average. That's because these computations can be, and very often are, thrown off dramatically by extreme values.

For those of you who see yourselves as quants, think in terms of "kurtosis," a measure of the thickness of the tails of a distribution. For some ratios, especially growth rates and valuation metrics, the levels of kurtosis are spectacularly high. The kurtosis of a normal distribution is about 3.00. (In Excel, you'd see this as zero, since the Excel KURT function gives you actual Kurtosis minus three, or "excess kurtosis." For the kinds of data sets we use, excess kurtosis can be above 100.

If you aren't a quant, think of it this way. A good growth rate might be around 20%. A great growth rate might be 30%. A stupendous growth rate might be 50%. How might you feel about a 7,000% growth rate? Believe it or not, we have lots of number like that, and unfortunately, they are just as likely to occur among very large companies as well as very small. When we get big companies with big kurtosis, we wind up with cap weighted averages that make no sense for investors.

When we object to such extremes, it doesn't mean we choose to wear blinders and ignore exceptional financials. Companies that are exceptional, for better or worse, should be noted as such and their experiences ought to be factored into industry averages, since they are part of the collective experiences of the industries in which they operate. But in truth, the extreme numbers we see in the database do not represent exceptional instances of business performance. They generally represent unusual events recorded by unusual accounting entries, the exact kinds of things security analysts are trained to eliminate from the analysis. (And by the way, the weeding out of unusuals was not invented by aggressive sell siders looking to polish numbers back in 2000; it's given vehement and conspicuous attention in the Graham & Dodd classic on Security Analysis. So when we look to cleanse our averages of extremes, we are not blinding ourselves to exceptional corporate performance; we are instead choosing to "blind ourselves" to numbers Graham

& Dodd urged analysts to ignore precisely because they distract us from, rather than represent, a company's true financial state of affairs.

Undertaking such a task through an objective algorithm applied automatically to a data set is, necessarily, an approximation. But we believe the method we're using is quite reasonable.

Our goal is to look to the central point in the distribution but temper the influence of very small companies by adding an element of weighting. One might think of what we're doing as a "thick weighted median."

Here's how our logic works.

Preliminary Dataset Trim

Eliminate ADRs
Eliminate companies with missing data
Eliminate companies for which the data-point is NA

Second Trim

Eliminate all companies with zero values when industry average is being calculated for the following factors

DIVIDEND INFORMATION: Yield5YAvg
DIVIDEND INFORMATION: Yield
FINANCIAL STRENGTH RATIOS: Payout5YAvg
FINANCIAL STRENGTH RATIOS: PayRatioTTM
FINANCIAL STRENGTH RATIOS: IntCovTTM
GROWTH RATES: Div%ChgA
GROWTH RATES: Div3YCGr%
GROWTH RATES: Div5YCGr%
GROWTH RATES: Retn%TTM

(Note: We are not eliminating zero values form the calculation of industry average debt ratios. When comparing capital structures, zero is a meaningful value for debt and/or equity. But when assessing payout ratios (an indication of how burdensome the dividend is), zero is not meaningful as an indicator of strength when the company pays no dividends.)

Third Trim

Starting with the sample that remains after the first two trims have been completed, we sort the group and eliminate the top 16.5% of the group, and the bottom 16.5%. This, in effect, means we're eliminating one stand deviation's

worth of observations, half from among the high values and half form the low values. Note, though, that we use these theoretical estimates of what percent of a data-set is captured by one stand deviation. We don't rigorously calculate and use the actual stand deviations for each data set since these are influenced by the extreme values we want to eliminate.

After the third trim, perform the final calculation:

If the number of companies remaining after the third trim is less than 10, then use the actual median

Otherwise (as is the case for most industries), use a market cap weighted average