ORIGINAL RESEARCH PAPER | Statisitics | Volume : 6 | Issue : 8 | August 2016 | ISSN - 2249-555X | IF : 3.919 | IC Value : 74.50

Statistical Models Applied to A Log-Normally Distributed Asset/Stock-Pricing Statistic: the Asset Gross Revenue Multiplier (AGRM)

KEYWORDS

Herman Kelting	Sadri Khalessi
PhD, PO Box 370820, Las Vegas, NV 89137	PhD, 893 University Avenue, Los Altos, CA 94024

ABSTRACT *The purpose of this paper is to utilize statistical tests to show the AGRM, a cross sectional price statistic, is a valid asset and stock pricing model based upon its distributions for nine SIC groups totaling 4,843 companies. The AGRM is the quotient of total assets at market capitalization value and total revenue. The importance of the AGRM is that it may be used to price total assets and common equity from total revenue.*

Popular statistical tests of the AGRM frequency distributions indicate that it is log-normally distributed for all nine SIC groups. The Box and Cox transformation indicate that the logarithmic transformation is appropriate to describe the distribution of the AGRM across the nine data sets. Bartlett's test for equality of variances showed that seven of the nine distributions of LNAGRM are highly similar and the remaining two are somewhat similar to the other seven. The standard deviations of the seven groups of LnAGRM vary from 0.66 to 0.78 and all nine vary from 0.66 to 0.99 with a coefficient of variation of 0.14.

Introduction

The purpose of this research is to provide statistical support for an asset and stock-pricing model using the Asset Gross Revenue Multiplier (AGRM), which may explain asset and stock prices. The AGRM has been subject of only one prior study by one of the authors. [Kelting [1]]

The AGRM is defined as

$$\text{AGRM} = \frac{\text{Total assets at market capitalization value (TAMCV)}}{\text{Total Revenue (TR)}}$$

TAMCV is the sum of the market capitalization value of common equity [CEMCV; Number shares CE outstanding (NSO) x Price per share] plus the cost of total liabilities including minority interests (TL). Cost rather than market value of total liabilities is used to have a readily available measure of TL amount from the financial statements and in computerized databases.

The AGRM is a cross sectional price statistic that is derived from market data and may be used to value total assets (*i.e.*, TAMCV = AGRM x TR) and common equity (*i.e.*, [CEMCV/share = (AGRM x TR – TL)/NSO] assuming no preferred stock. The AGRM implicitly takes earnings predictions and the discount rate into account, while the discounted cash flow model (*i.e.*, DCFM; $PV = \sum Rt/(1+k)^t$] makes explicit predictions of these in Rt, annual cash flow, and k, the discount rate often defined by the weighted average cost of capital. Indeed, the DCFM must be used for project investments to assure that financial management has support for the time dimensions of revenue, costs, and expenses for a new project that contributes to the NPV of the firm. Exogenous explicit predictions of these by investment analysts for asset and stock pricing are supremely difficult because of their inadequate knowledge of business operations.

The nearest cross sectional price statistic to the AGRM is the Enterprise Value to Sales (EVS) ratio. Most frequently, the EVS ratio is defined as the sum of common and preferred equity plus interest bearing debt less cash divided by sales or total revenue; thus, both cash and current liabilities are omitted in enterprise value. But there are slight variations in the definition of enterprise value [Alexson [2]; Bhojraj [3]; Casta [4]; Compustat [5]; Harbula [6]; Kim [7]; Lie [8]; Panteleo [9]; White [10]].

The advantages of the AGRM to value total assets or common equity are that it (1) matches the highest level of total assets and revenue based upon the accounting identity, TA = TL + Equity Capital from the financial statements, (2) takes into account all operating financial requirements necessary for business operations, including project investments, acquisitions, and extra cash required for expected downturns in business activity, (3) is simple to calculate knowing only TAMCV and total revenue, (4) and may be used by market participants to set stock prices based upon the frequency distributions in this study across nine standard industrial classifications (SICs) and the earlier multiple regression analysis by one of the authors.

Materials and Methods

The database for AGRM research

In order to determine if stock prices in the public markets are set by the AGRM, we formulated a research plan to calculate the AGRM by industry group and used Standard and Poor's Capital IQ database at the Graduate School of Business Library at Stanford University. This database contained about 9,300 lines of data (one line per company). After elimination of cases containing missing data or outliers, we had 6,039 AGRM-calculated companies. We grouped these companies by four-digit Standard Industrial Classification ("4DSIC").

Preliminary exploratory analysis of the AGRM data led us to hypothesize that the distributions of the AGRMs would be log normal with a long right tail. Furthermore, large AGRMs (*e.g.*, 500) would skew the distributions of the four-digit SICs (4DSIC). Thus, we used the Coefficient of Variation, CV<1, as the criteria to eliminate some of these outliers; however, the CV<1 criterion was avoided if the distribution of large AGRMs was continuous in the sense of a large number of continuously large AGRMs. This set of criteria led to 5.3% reduction in the total number of com-

Volume : 6 | Issue : 8 | August 2016 | ISSN - 2249-555X | IF : 3.919 | IC Value : 74.50

panies considered. We originally predicted five-percent outliers; thus the 6,039 companies for which we calculated the AGRM, was reduced by 318 outliers (5.3 percent) to a net number of companies of 5,760. We did not calculate the cumulative percentage of 4DSIC outliers to assure independence between 4DSIC group outlier eliminations. We further eliminated 917 companies in ten four- digit SICs because their AGRMs differed substantially from the remaining companies in each of the nine, four-digit SIC groups yielding 4,843 companies in nine, four-digit SIC groups subject to distributional analysis.

An introduction to statistical analysis of the AGRM
In statistical literature the terminology "modeling" applies to two distinct concepts. Namely, curve fitting and distribution fitting. These two concepts are very different and require different kinds of data analysis. One is fitting a curve to a set of points, which involves modeling a response variable as a function of one or more covariates or predictor variables. Simple and multiple regression, general linear and none-linear modeling, Ridge regression, The Lasso, and time series analysis are appropriate in this case with a wide range of applications. They are part of the so-called supervised learning as discussed by Hastie et.al [11].

Distribution fitting on the other hand falls under the umbrella of unsupervised learning. We are not required to make any prediction in this case. It involves modeling the probability distribution of one or more variables. In the univariate case one seeks to model the probability density function of a variable. The model usually involves one or more parameters. In the multivariate case the joint probability distribution is often the goal of the researcher. In either case the model is a normalized probability density function. There are several software packages, which fit distributions to data interactively. One such package is:

UNIFIT [12], which is an interactive computer package for fitting probability distributions to observed data.

Our objective in this paper is to find, fit, and statistically test a univariate family of distributions, which best describe the distribution of each of the nine AGRM data groups mentioned earlier. We first do some exploratory analysis by computing the usual sample moments, namely, mean, standard deviation, coefficient of skewness, and coefficient of kurtosis, as well as histograms of each of the nine groups of data. Based on this descriptive analysis, we postulate the likely distributions, which might describe the data adequately. Finally, we perform various statistical tests on each data set to assess the goodness of the fitted distribution. In particular we consider the lognormal as a viable model for fitting the available data from the nine industry segments.

Exploratory analysis and descriptive statistics
As previously stated, we will demonstrate that AGRM is a simple and viable multiplier in the asset pricing process and financial market analysis. The available data consist of the AGRM multiplier for 4843 companies. These companies are from nine different market segments. Table-1 and Table-2 depict the descriptive statistics for both the AGRM and Ln-AGRM, the natural logarithm of AGRM respectively for all nine data sets.

Table-1

VAR →	AGRM 1	AGRM 2	AGRM 3	AGRM 4	AGRM 5	AGRM 6	AGRM 7	AGRM 8	AGRM 9
N (sample size)	415	157	477	1253	537	442	753	649	160
Mean	6.5	2.1	1.6	2	3.2	1.1	9.3	3.3	1.7
Median	4.9	1.8	1.2	1.4	2.7	.74	7	2.4	1.3
SD	5.7	1.5	1.3	1.8	2.5	1.1	9	3.1	1.2
Min	.39	.31	.22	.25	.11	.1	.1	.05	.25
Max	47.83	8.75	7.93	18.51	16.79	9.61	66.95	21.49	7.77
Range	47.44	8.44	7.71	18.26	16.68	9.51	66.84	21.44	7.52
Skewness	3.17	1.82	2.1	3.6	2.1	3.7	2.5	2.3	1.93
Kurtosis	17.16	7.2	8.35	23.04	9.22	22.76	11.7	10.11	7.93
SE	.28	.12	.06	.05	.05	.05	.33	.12	.07

Table-2

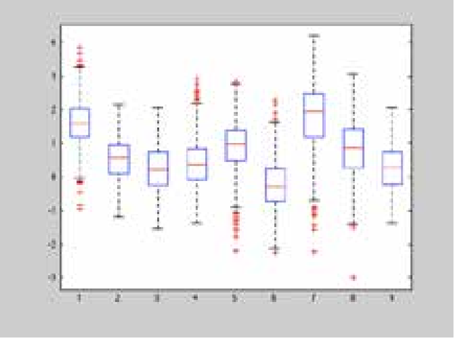
VAR →	Ln-AGRM1	Ln-AGRM2	Ln-AGRM3	Ln-AGRM4	Ln-AGRM5	Ln AGRM6	Ln AGRM7	Ln AGRM8	Ln AGRM9
N(sample size)	415	157	477	1253	537	442	753	649	160
Mean	1.6	.53	.23	.4	.9	-.24	1.81	.85	.27
Median	1.6	.57	.21	.36	.99	-.3	1.95	.86	.29
SD	.72	.66	.7	.71	.78	.74	.99	.86	.69
Min	-.94	-1.17	-1.51	-1.39	-2.21	-2.3	-2.21	-3	-1.39
Max	3.87	2.17	2.07	2.92	2.82	2.26	4.2	3.07	2.05
Range	4.81	3.34	3.58	4.3	5.03	4.57	6.41	6.07	3.44
Skew	-.06	-.07	.09	. 4	-.54	.34	-.56	-.15	-.04
Kurt	3.8	2.71	2.76	3.12	3.78	3.31	3.53	3.18	2.7
SE	.04	.05	.03	.02	.03	.04	.04	.03	.05

Volume : 6 | Issue : 8 | August 2016 | ISSN - 2249-555X | IF : 3.919 | IC Value : 74.50

Notes and observations:

1. The mean of AGRM is larger than its median for all nine groups, indicating a positively skewed distribution.
2. The mean and median are essentially the same for the Ln-AGRM for all nine groups, indicating a symmetric distribution.
3. The SD's of all nine groups for the Ln-AGRM are fairly similar varying from 0.66 to 0.99 (mean = 0.76, SD = .10, CV = 0.14), meaning that the market dynamics for all nine industry segments and their volatility as measured by Ln- AGRM is relatively stable.
4. The means of Ln-AGRM are different for the nine market segments, as each industry has it's particular characteristic.
5. Kurtosis coefficient is 3 in case of the normal distribution. The estimated kurtosis range from 2.7 to 3.8, which is not very different from 3 for the normal distribution.
6. Finally, the coefficient of Skewness of Ln-AGRM is less than 0.57 in absolute value for all groups. This coefficient is zero in case of the normal distribution. Based on the above observations we suspect that lognormal distribution might be a viable candidate to describe the AGRM distribution. We shall perform various rigorous test of normality of Ln-AGRM for all nine groups. Figure-1 depicts the boxplot of Ln-AGRM for all nine datasets. The corresponding histograms are given in Figure-2.

Figure-1, Ln-AGRM Box Plots



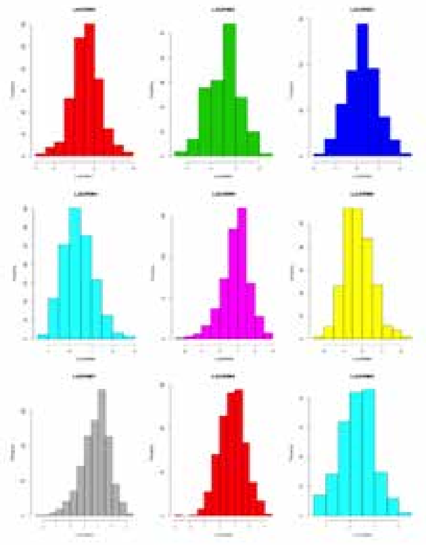


Figure-2 Histogram of Ln-AGRM for all nine datasets

Distribution fitting and parameter estimation

We explored several positively skewed distributions as possible candidates to our data. See Johnson et.al [13]. For brevity, we only report the analysis for the AGRM1, the data set from the first market segment. The analysis led us to conclude that the lognormal describes the AGRM distribution well. A random variable X is said to have a lognormal distribution if the logarithm X has a normal distribution. The mean and variance of the log normal distribution are functions of the mean and variance of the corresponding normal distribution as the following formulae indicates.

$$X \sim \text{log normal}(\mu,\sigma) \Rightarrow \text{Mean}(X) = \mu,\ \text{sd}(X) = \sigma$$

$$Y = \log(X) \Rightarrow Y \sim \text{log normal}(m,s)$$

$$m = \ln\left(\frac{\mu^2}{\sqrt{\sigma^2+\mu^2}}\right),\quad s = \sqrt{\ln\left(1+\frac{\sigma^2}{\mu^2}\right)}$$

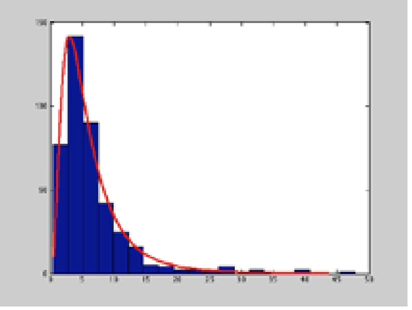


Figure-3: Distribution Function of AGRM1 & the fitted lognormal

The difference between normal and lognormal variability is that both forms of variability are based on a variety of forces acting independently of one another. A major difference, however, is that the effects can be additive or multiplicative, leading to normal or lognormal distributions, respectively. This property is particularly interesting and applicable in finance where, a variable might be modeled as lognormal if it can be thought of as the product of many independent positive random variables. This is the direct consequence of the central limit theorem in the log-space. In finance, we often use the compound return from a sequence of many independent returns (each expressed as its return +1); or long-term discount factor can be derived from the product of short-term discount factors. Considering the AGRM1 in log scale therefore seems to be reasonable in this finance application when dealing with discounted factors or compound interest that AGRM may inherently represent.

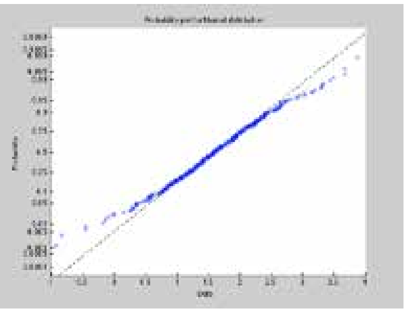


Figure-4 Normal Probability Plot of Ln-AGRM1 Slightly skewed to the left with negative coefficient of skewness of -0.06

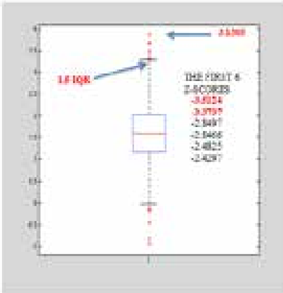


Figure-5 Boxplot of Ln-AGRM1: The distribution is more peaked than normal with Kurtosis of 3.8 compared to 3 for normal and it is slightly skewed to the left with Skewness of -.06 compared to 0 for normal. Only 8 observations out of 415 are outside the +/- 3-sigma limits. We conclude that Ln-normal is a reasonable fit for the AGRM1 data set.

Fitting alternate distributions to AGRM1

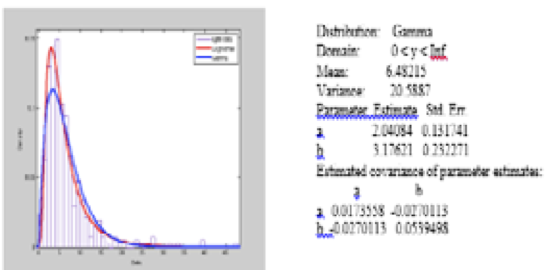


Figure-6 Gamma Distribution and the corresponding Maximum likelihood parameter estimates

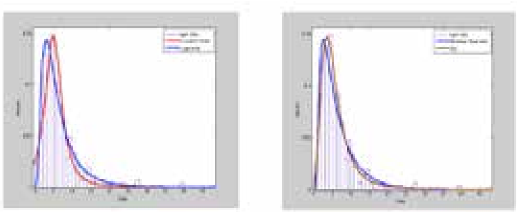
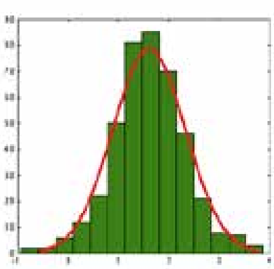


Figure-7: Fitting t-location-Scale, lognormal, Birnbaum-Saunders, and Bur distribution to AGRM1.

Figure-7: Fitting t-location-Scale, lognormal, Birnbaum-Saunders, and Bur distribution to AGRM1.



Mean (Ln-AGRM1)= 1.6045, STD (Ln-AGRM1)= 0.7221

95% Confidence interval for the mean and SD

Mu (1.5347 1.6769), sigma (0.6769 .7759)

FIGURE-8 The Ln-AGRM1 empirical distribution, the fitted normal curve, the estimated mean and standard deviation and the corresponding 95% confidence interval

Fitting lognormal distribution to AGRM for all nine market-segments.

We performed similar analysis for the other eight data sets. The lognormal proved to be a reasonable model for those data sets as well. In this section we show each fitted lognormal distribution and the corresponding parameter estimates as well as the 95% confidence interval for the mean and standard deviation. The graphs are generated using the Mat-lab (Math-Works) *dfittool*. We also used the *kdensity* tool of Mat-lab with the normal kernel to estimate the distribution none-parametrically. However, the resulting fitted distributions were not significantly different than the parametrically fitted lognormal distributions.

AGRM1 Lognormal distribution

Mu = 1.60446 [1.5347, 1.67423]

Sigma = 0.723007 [0.676938, 0.775857]

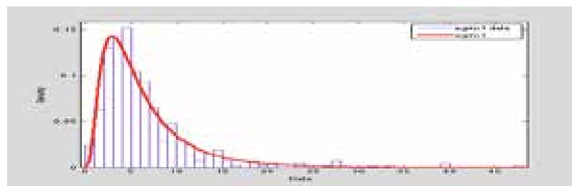


Figure-9 fitted lognormal distribution to AGRM1: N=415

AGRM2 Lognormal distribution

Mu = 0.53022 [0.426943, 0.633497]

Sigma = 0.655121 [0.589791, 0.736855]

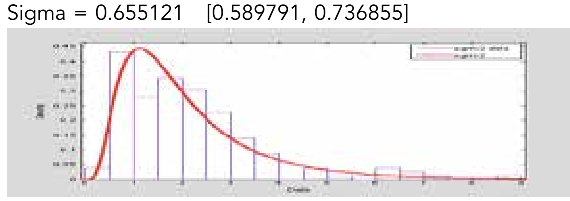


Figure-10 fitted lognormal distribution to AGRM2: N=157

AGRM3 Lognormal distribution

Mu = 0.227172 [0.164114, 0.29023]

Sigma = 0.700885 [0.659052, 0.748432]

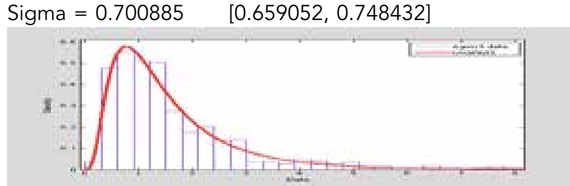


Figure-11 fitted lognormal distribution to AGRM3: N=477

AGRM4 Lognormal distribution

Mu = 0.401136 [0.361966, 0.440307]
Sigma = 0.706744 [0.680116, 0.735558]

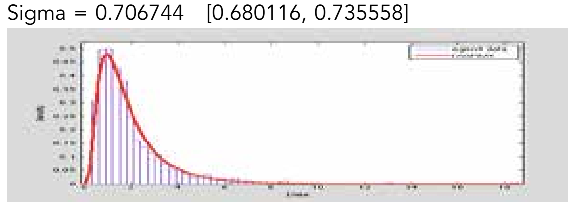


Figure-12 fitted lognormal distribution to AGRM4: N=1253

Volume : 6 | Issue : 8 | August 2016 | ISSN - 2249-555X | IF : 3.919 | IC Value : 74.50

AGRM5 Lognormal distribution
Mu = 0.899827 [0.834048, 0.965605]

Sigma = 0.775962 [0.732164, 0.825376]

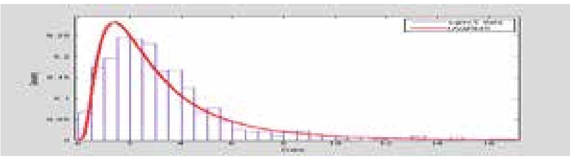


Figure-13 fitted lognormal distribution to AGRM5: N=537

AGRM6 Lognormal distribution
Mu = -0.242647 [-0.311721, -0.173573]

Sigma = 0.738896 [0.693186, 0.791109]

Figure-14 fitted lognormal distribution to AGRM6: N=442

AGRM7 Lognormal distribution
Mu = 1.81382 [1.74316, 1.88449]

Sigma = 0.987791 [0.940296, 1.04038]

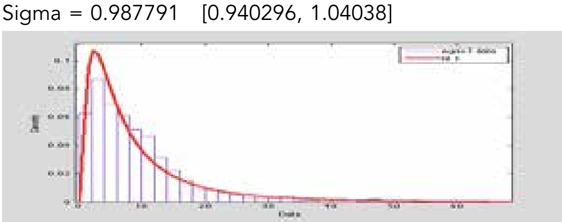


Figure-15 fitted lognormal distribution to AGRM7: N=753

AGRM8 Lognormal distribution
Mu = 0.844846 [0.778522, 0.91117]

Sigma = 0.860467 [0.816065, 0.910017]

Figure-16 fitted lognormal distribution to AGRM8: N=649

AGRM9 Lognormal distribution
Mu = 0.272664 [0.164804, 0.380524]

Sigma = 0.690801 [0.6225, 0.77607]

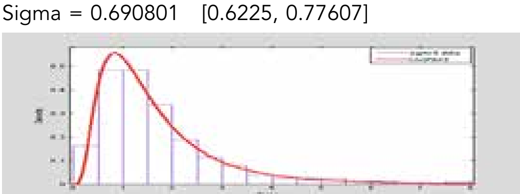


Figure-17 fitted lognormal distribution to AGRM9: N=160

The Box and Cox transformation of the data
In this section we further transform the AGRM in each group using the Box and Cox power transformation [14], in search of a function of AGRM, which might describe the data more accurately.

The Box–Cox power transformation is given by:

$$y_i^{(\lambda)} = \begin{cases} \dfrac{y_i^\lambda - 1}{\lambda}, & \text{if } \lambda \neq 0, \\ \log(y_i), & \text{if } \lambda = 0. \end{cases}$$

Such transformation is often used to stabilize the variance and reduce skewness in the data distribution, thus making the data more normal. One application is the general linear models where the errors are assumed to be independent and normally distributed with equal variances.

The maximum likelihood estimate of λ-ηατ for which the sum of squares of error is minimum is obtained computationally. If the optimal value of λ is zero, the log transformation is the most appropriate one. However, if λ-hat is other than zero, the power transformation given in the above formula is better. For groups 4-7 the λ-hat values differ from zero, but even in these cases the departure from zero is not substantial, indicating the transformation of the AGRM is the appropriate one as indicated in Table 3.

AGRM	λ-hat	Ln	AGRM	λ-hat	Ln
AGRM1	0.0	-	AGRM6	-0.1	x
AGRM2	0.0	-	AGRM7	0.17	x
AGRM3	0.0	-	AGRM8	0.0	-
AGRM4	-0.2	x	AGRM9	0.0	-
AGRM5	0.2	x			

Table-3 λ-hat values

Statistical test for normality
There are many statistical tests for normality. Some are more sensitive than others. Our Null hypothesis **H_0:** states that the distribution of the $AGRM_i$, i=1,2…9; follow the normal distribution vs. the alternative hypothesis **H_1:** It follows a non-normal distribution**.** These tests are done separately for each data set**.** By definition, the p-value for each is the smallest level of significance (the α value or the type-I error) that the null hypothesis may be rejected.

One popular test is the Kolmogorov Smirnov test [15]…[21]; of normality on the Ln-AGRM as well as the Box & Cox power transformed data sets. Table-4 shows the P-values for these tests. If the value in the **H** Colum is zero, there is no significant departure from normality. If the value is one, there is statistical evidence of significant departure from normal distribution. Note that the values with a 1 in the **H** column correspond to p-values less that α=0.05.

In table-4, the P-values are all greater than 0.05, after the Box and Cox transformation of the data.

H	P-Val Ln-transformation	Transformation	P-Val Box & Cox
0	.58	Ln	.58
0	.53	Ln	.53

Volume : 6 | Issue : 8 | August 2016 | ISSN - 2249-555X | IF : 3.919 | IC Value : 74.50

0	.71	Ln	.71
1	.03	-((X^(-.2))-1)/.2	.90
1	.02	((X^.2)-1)/.2	.57
0	.48	-((X^.1)-1)/.1	.95
1	.01	((X^.17)-1)/.17	.47
0	.99	Ln	.99
0	.85	Ln	.85

Table-4: Kolmogorov Smirnov Test for normality of Ln-data

LAGRM4, LAGRM5, and LAGRM7 do not pass the KS normality test. We therefore perform the Box & Cox power transformation on these variables. After this transformation, they all pass the KS test for normality at α=0.05 significant level, as indicated by the p-values in the last column. If the p-value is less than α=0.05, we reject the normality assumption. This is the case for AGRM4, AGRM5, and AGRM7, indicated by red color in table-3. Note that for AGRM6, the value of **λ-hat** is 0.1, which is the closest to 0. In this case the Ln-transformation seems adequate and the KS test for normality does not reject H_0. This analysis shows that the Box and Cox transformation actually achieves near normal distribution for the AGRM in the nine industry segments.

Other statistical tests for normality
Chi-SQ goodness of fit test, [22]…[23]:

The chi-square test for normality uses the statistics:

$$\chi_r^2 = \Sigma \frac{(\text{observed} - \text{expected})^2}{\text{expected}}.$$

The data range is devided into several non-overlapping intervals and the frequency of the data in each interval is compared with what would be expected if the data followed a normal distribution. If the value of test statistics χ_r^2 is large as compared to the α-percentile of the χ_r^2 distribution where r is the degrees of freedom, we reject the normality assumption. Here r is number of intervals considered.

After the Ln-transformation the data has been further normalized and are tested against the standard normal distribution. The corresponding variables after normalization are: z_1-z_9. Table-5 depicts the results, obtained using the Mat-lab Chi2-GOFFIT tool.

z=(Ln-AGRM-Mean(Ln-AGRM))/SD(Ln-AGRM)	H	P-Value
1	0	0.09
2	0	0.06
3	0	0.32
4	0	0.85
5	0	0.11
6	0	0.49
7	0	0.06
8	0	0.95
9	0	0.60

Table-5: Chi-Square for normality

The fact that H=0 in table-5, indicates we have no statistical evidence that the standardized transformed data do not differ significantly from the standard normal distribution at α= 0.05. Note that the p-values are all greater than .05

The Lilliefors Test of normality
We further perform the Lilliefors test [24]…[29], which is a two-sided goodness-of-fit test suitable when the parameters of the null distribution are unknown and must be estimated. This is in contrast to the one-sample Kolmogorov-Smirnov test, which requires the null distribution to be completely specified.

$$D = \max_x |F(x-) - G(x)|.$$

The Lilliefors test statistic is Where F (x) is the empirical cdf of the sample data and G is the cdf of the hypothesized distribution with estimated parameters equal to the sample parameters (In this case the normal distribution).

Table-6 summarizes the results, again, indicating no significant departure of the transformed data from the normal distribution. The Lillie test tool in Mat-lab was used to get these results. Each z, is the normalized Ln-AGRM data and **α=0.01**

z=(Ln-AGRM-Mean(Ln-AGRM))/SD(Ln-AGRM)	H	P-Value
1	0	0.17
2	0	0.13
3	0	0.30
4	0	0.50
5	0	0.16
6	0	0.50
7	0	0.09
8	0	0.50
9	0	0.50

Table-6 Lilliefors test of normality.

The Shapiro-Wilk Test (R-Software)
Finally we perform the Shapiro-Wilk test, [30]…[32], which again confirms the previous conclusions of no significant departure of the transformed data from the normal distribution with α=. 05. Results are given in Table-7

Z=(Ln-AGRM-Mean(Ln-AGRM))/SD(Ln-AGRM)	H	P-Value
1	0	0.02
2	0	0.29
3	0	0.16
4	1	0.65
5	1	0.99
6	0	0.57
7	1	0.17
8	0	0.14
9	0	0.74

Table-7 Shapiro-Wilk Test of normality

Bartlett's test for equality of variances
Based on the data in tables 1 and 2 of section II; we stated that the variances of the Ln-AGRM are very similar for all nine groups. Six of the nine data sets have standard deviations between 0.66 and 0.74 and all are in the range of 0.66 to 0.99 with average SD of 0.76, standard devia-

Volume : 6 | Issue : 8 | August 2016 | ISSN - 2249-555X | IF : 3.919 | IC Value : 74.50

tion of 0.10, and coefficient of variation (CV) of 0.14. This seems to be the smallest CV for a business process and suggests the stock price component of the AGRM is set by the efficient-market hypothesis with common price determinants across all nine groups.

We now perform a test known as the Bartlett's test [33], to verify similarity in the variances of Ln-AGRM statistically. The Bartlett's statistics is 164 with 8 degrees of freedom. With this large value of the statistic, the p-value is very small, indicating the rejection of the hypothesis of equality of the variances of all nine groups. Further examination of the estimate of the standard deviations indicate that groups 7 and 8 have larger variance than the other 7 groups. Excluding these two groups and performing the Bartlett's test again, we get the Bartlett's statistics of 11.6, with 6 degrees of freedom and a p-value of 0.07, which is not significant at $\alpha=0.05$. We therefore conclude that the standard deviations of AGRM are statistically the same for groups 1-6, and 9. A third iteration of Bartlett's test for group 7 and 8 yielded a statistic value of 13.1 with one degree of freedom and a p-value of close to zero, indicating that the standard deviations of group 7 and 8 are statistically different. Although, these statistical tests indicate that the true standard deviation of AGRM are not all the same, from the practical point of view however, they are very comparable. This conjecture is further verified by looking at the 95% confidence intervals for each group standard deviation given in figures 10-17.

Fitting distributions to AGRM using the Pearson System.

The statistician Karl Pearson devised a system, or family, of distributions that includes a unique distribution corresponding to every valid combination of mean, standard deviation, skewness, and kurtosis [34]…[36], Given sample values for each of these moments from data, it is easy to find the distribution in the Pearson system that matches these four moments and to generate a random sample of that type of distribution. The Pearson system includes seven basic types of distribution together in a single parametric framework. It includes common distributions such as the normal and t distributions, simple transformations of standard distributions such as a shifted and scaled beta distribution and the inverse gamma distribution, and one distribution—the Type IV—that is not a simple transformation of any standard distribution. *For a given set of moments, there are distributions that are not in the system that also have those same first four moments, and the distribution in the Pearson system may not be a good match to the data, particularly if the data are multimodal.* But the system does cover a wide range of distribution shapes, including both symmetric and skewed distributions. We used Mat-lab (Pearsrnd), to compute the first four moments and used them as input. The output of Pearsrnd provides the type and the coefficients of the distribution in the corresponding system. The functions also generate random numbers from that specific distribution.

The seven distribution types in the Pearson system correspond to the following distributions:

Normal Distribution

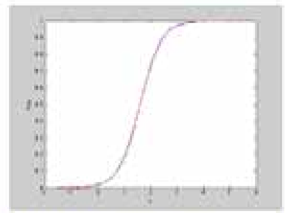
1. Four-parameter generalized beta distribution
2. Symmetric four-parameter beta distribution
3. Three-parameter gamma distribution
4. Not related to any standard distribution. The density is proportional to:

$$(1 + ((x - a)/b)^2)^{-c} \exp(-d \arctan((x - a)/b))$$

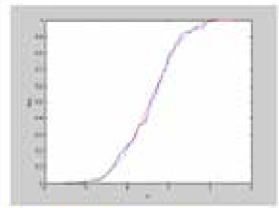
5. Inverse gamma location-scale distribution
6. F location-scale distribution
7. Student t- location scale distribution

We thus obtained the Pearson-type distribution for each of the nine, Ln-AGRM and the nine original AGRM based on the first four sample moments. The fitted CDF are given below. The moments are given in Table-2 of section II.

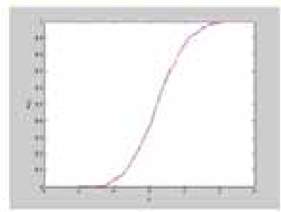
Ln-AGRM1, Pearson Type-4



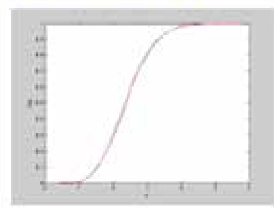
Ln-AGRM2, Pearson Type-1



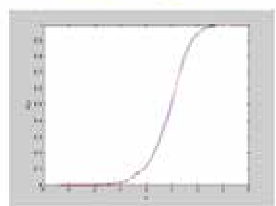
Ln-AGRM3, Pearson Type-1



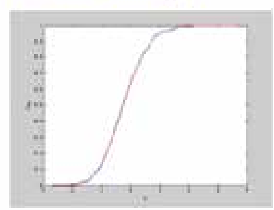
Ln-AGRM4, Pearson Type-1



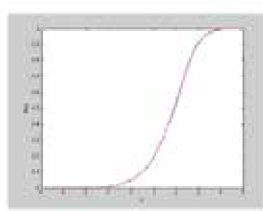
Ln-AGRM5, Pearson Type-4



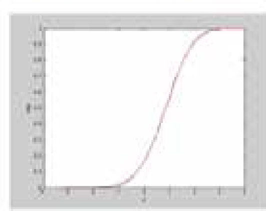
Ln-AGRM6, Pearson Type-4



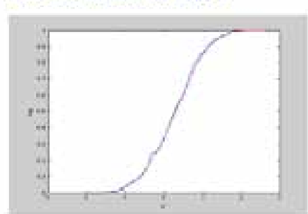
Ln-AGRM7, Pearson Type-6



Ln-AGRM8, Pearson Type-4



Ln-AGRM9, Pearson Type-1



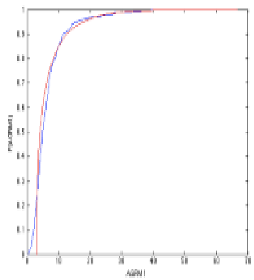
Summary of Pearson types for Ln AGRM:

Ln-AGRM#	Mean	SD	Skew-ness	b_1=Skewness^2	b_2=Kurtosis	Type
1	1.6	0.72	-0.06	0.0036	3.8	IV
2	0.53	0.66	-0.07	0.0049	2.71	I
3	0.23	0.7	0.09	0.0081	2.76	I
4	0.4	0.71	0.04	0.0016	3.13	I
5	.90	0.78	-0.54	0.2916	3.78	IV
6	-.24	0.74	0.33	0.1089	3.31	IV
7	1.81	0.99	-0.56	0.3136	3.51	VI
8	0.85	0.86	-0.15	0.0225	3.18	IV
9	0.27	0.69	-0.04	0.0016	2.7	I

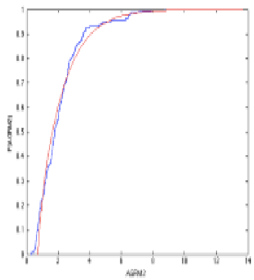
Volume : 6 | Issue : 8 | August 2016 | ISSN - 2249-555X | IF : 3.919 | IC Value : 74.50

We also performed the analysis for AGRM and the analysis is given below

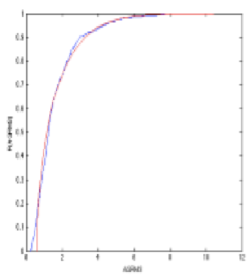
AGRM1, Pearson Type-1



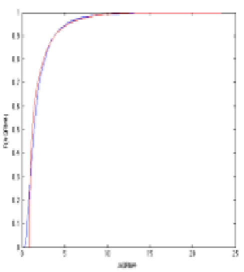
AGRM2, Pearson Type-1



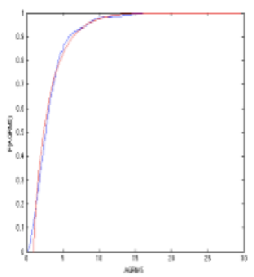
AGRM3, Pearson Type-1



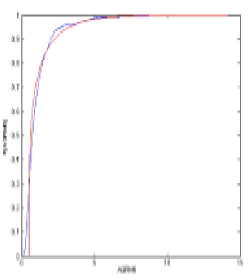
AGRM4, Pearson Type-VI



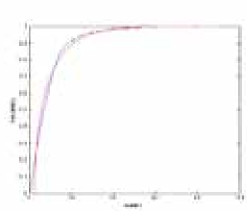
AGRM5, Pearson Type-1



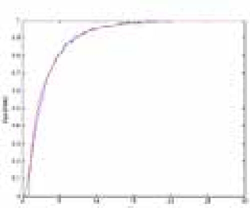
AGRM6, Pearson Type-1



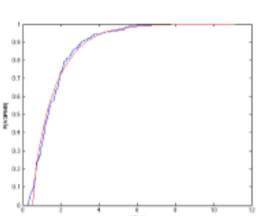
AGRM7, Pearson Type-1



AGRM8, Pearson Type-1



AGRM9, Pearson Type-1



Summary of Pearson types for AGRM:

AGRM#	Mean	SD	b_1=Skewness^2	Skew-ness^2	Kurto-sis	Type
1	6.45	5.69	3.2	10.24	17.2	I
2	2.1	1.45	1.82	3.31	7.2	I
3	1.6	1.26	2.07	4.41	8.35	I
4	1.96	8.34	3.6	12.96	23.04	I
5	3.22	2.48	2.1	4.41	9.2	I
6	1.06	1.05	3.7	13.69	22.76	IV
7	9.31	9.01	2.9	6.25	11.7	I
8	3.32	3.1	2.3	5.29	10.1	I
9	1.66	1.2	1.93	3.73	7.92	I

Based on this analysis, Pearson type-1 seems to be the best candidate for the AGRM. This family of distributions is the so called generalized Beta (GB) distribution, which is a continuous probability distribution with five parameters, including more than thirty named distributions as limiting or special cases. It has been used in the modeling of income distribution, stock returns, as well as in regression analysis. The exponential generalized Beta (EGB) distribution follows directly from the GB and generalizes other common distributions, [37]. Normal distribution is the limiting distribution for Pearson types: I, III, IV, V, and VI.

Results and Discussions

In this paper, we perform several different tests for normality, namely, the usual qq-plot, the nonparametric Chi2-Goodness of Fit test, the Kolmogorov Smirnov test, Lilliefors test and the Shapiro Wilkes test. The test results consistently show the effectiveness of the logarithmic transformations performed on AGRM, when we test the null hypothesis H_0: that each transformed variable follows a normal distribution vs. an alternative H_1: It follows a non-normal one.

For each variable, the Basic methodology is to first create a proper distribution function from the empirical sample distribution, non-parametrically, that is without assuming any particular parametric distribution a priori; and test it against the normal distribution. The parametric approach is to assume lognormal distribution a priori and estimate it's parameters using the maximum likelihood method. The normality assumption is then checked using the goodness of fit test. For brevity we only report the test for normality in the parametric case. The results were similar for the empirical distributions created form sample data.

The first data set, AGRM1, was initially used as a training set to establish the methodology. This data set consisted of 415 observations in metal mining, crude petroleum, natural gas, etc. Several analytic approaches were adopted. These included obtaining basic descriptive statistics and various moments as well as plotting the empirical PDF (Histogram), and CDF of the data. The initial plot of the empirical distribution revealed a positively skewed distribution suggesting the lognormal distribution as a possible candidate for the population distribution of AGRM1. However we explored other right skewed distributions such as Gamma, Burr, Birnbaum-Saunders, and t-location scale in addition to the lognormal distribution. The graphs for the corresponding fitted curves for these distributions are given in section II.

This preliminary analysis revealed that the lognormal distribution was indeed the best candidate for the data at hand.

Volume : 6 | Issue : 8 | August 2016 | ISSN - 2249-555X | IF : 3.919 | IC Value : 74.50

The lognormal had the additional advantage of tractability, having the desirable properties of the Normal distribution, once the data is transformed to the log scale. The log transformation is to the base e and is denoted by Ln.

The next step was to fit the lognormal distribution to all nine sets of data. We did this using various tools in Matlab including the dfittool, distfit, makedist, ksdensity, and mle tools. We also used the open source R software for parts of the analysis where the Matlab tools were not readily available or cumbersome to use.

We further applied the Box and Cox transformation, which confirmed the appropriateness of the lognormal model for some of the nine data sets. The computed maximum likelihood estimate of λ, the parameter used in the Box and Cox method, applied to the nine sets are remarkably close to zero, indicating that the log transformation of the AGRM data may indeed be the appropriate model for achieving normality. In other words, Ln-AGRM is nearly normal. Additionally we analyzed each data set using the Pearson System of distributions based on the first four sample moments. The results indicated the generalized type-1 β-distribution might be the best parametric model for AGRM. Pearson system distribution is a very rich family of distributions and is often applied in financial analysis. See [38]…[40]

Conclusion

The significance of this paper is two-folded. First, it proposes the novel and yet simple cross-sectional multiplier, Asset Gross Revenue Multiplier (AGRM), to price total assets and common equity. The AGRM may be used to evaluate a company or a market segment in an effective and straightforward manner. Second, as far as we know, the in-depth and rigorous statistical treatment of the subject is new. The use of real market data to characterize the AGRM for nine market segments statistically and ascertaining that this multiplier follows a lognormal in most market segments and nearly lognormal in others is a new finding. This paper is a prelude to a subsequent paper in which we seek to find well-known common and simple market indicators covariates to estimate the AGRM using general linear models. The analysis performed here is significant, because it establishes the parametric distribution of AGRM to be used later as the dependent variables in the linear models for which the normality assumption is generally required. AGRM is a simple multiplier of total revenue to price assets and common equity. We hope to model AGRM, as a function common simple covariates in a follow up paper.

References

[1] Kelting, Herman. 2006. Multiple Regression Analysis and the Asset Gross Revenue Multiplier. *The Value Examiner,* (May/June): 29-32.

[2] Alexson, Ulf, Tim Jenkinson, and Per Stromberg. 2010. Borrow Cheap, Buy High? The Determinants of Leverage and Pricing in Buyouts. Working Paper 15952, *National Bureau of Economic Research*, (April): 1- 65.

[3] Bhojraj, Sanjeev and Charles M.C. Lee. 2002. Who Is My Peer? A Valuation-Based Approach to the Selection of Comparable Firms. *Journal of Accounting Research*. 10(2): 407-439.

[4] Casta, Jean-Francois, Luc Paugam, and Herve Stolowy. 2011.Non-additivity in accounting valuation: Internally generated goodwill as an aggregation of interacting assets. European Accounting Association 2011 annual meeting, Rome-Siena, Italy: 1-47.

[5] Compustat Data Glossary

[6] Harbula, Peter. 2009. Valuation Multiples: Accuracy and Drivers Evidence from the European Stock Market. *Business Valuation Review.* 28(4): 186-200.

[7] Kim, Moonchul, and Jay R. Ritter. 1999. Valuing IPOs. *Journal of Financial Economics.* 53: 409-437.

[8] Lie, Eric and Heidi J. Lie. 2003. Multiples Used to Estimate Corporate Value. *Business Valuation Digest.* 9(2): 1-13.

[9] Panteleo, Peter V. and Barry W. Ridings. 1996. Reorganization Value. *The Business Lawyer. 51 (February): 419-442.*

[10] White, James J. 2008. Bankruptcy Noir. *Michigan Law Review.* 106 (February 2008): 691-719.

[11] Trevor Hastie, Robert Tibshirani, and Jerome Friedman, The elements of Statistical Learning by 2nd edition

[12] http://dl.acm.org/citation.cfm?id=801536

[13] Johnson, N.L., S. Kotz, and N. Balakrishnan 1994 *Continuous Univariate Distributions, Volume 1*, (city: Wiley-Interscience,)

[14] Box, G. E. P. and Cox, D. R. 1964. An analysis of transformations (with discussion) *Journal of the Royal Statistical Society B*, 26: 211–252

[15] Z. W. Birnbaum and Fred H. Tingey. 1951. One-sided confidence contours for probability distribution functions. *The Annals of Mathematical Statistics*, 22(4): 592– 596

[16] William J. Conover. 1971. *Practical Nonparametric Statistics.* New York: John Wiley & Sons. 295–301 (one-sample Kolmogorov test), 309–314 (two-sample Smirnov test)

[17] Durbin, J. 1973. *Distribution theory for tests based on the sample distribution function*, SIAM

[18] George Marsaglia, Wai Wan Tsang and Jingbo Wang. 2003. Evaluating Kolmogorov's distribution. *Journal of Statistical Software*, 8(18)

[19] Massey, F. J. 1951. The Kolmogorov-Smirnov Test for Goodness of Fit, *Journal of the American Statistical Association* 46(253): 68–78

[20] Miller, L. H. 1956 Table of Percentage Points of Kolmogorov Statistics *Journal of the American Statistical Association* 51(273): 111–121

[21] Marsaglia, G., W. Tsang, and J. Wang. 2003. Evaluating Kolmogorov's Distribution. *Journal of Statistical Software* 8(18)

[22] Moore, D.S. 1986. Tests of the chi-squared type In: D'Agostino, R.B. and Stephens, M.A., eds.: Goodness-of-Fit Techniques. Marcel Dekker, New York

[23] Thode Jr., H.C. 2002. Testing for Normality. Marcel Dekker, New York

[24] Dallal, G.E. and Wilkinson, L. 1986. An analytic approximation to the distribution of Lilliefors' test for normality The American Statistician, 40: 294–296

[25] Stephens, M.A. 1974. EDF statistics for goodness of fit and some comparisons Journal of the American Statistical Association, 69: 730–737

[26] Thode Jr., H.C. 2002. Testing for Normality. Marcel Dekker, New York.

[27] Conover, W. J. 1980 *Practical Nonparametric Statistics*, Hoboken, NJ: John Wiley & Sons, Inc.

[28] Lilliefors, H. W. 1967. On the Kolmogorov-Smirnov test for normality with mean and variance unknown *Journal of the American Statistical Association* 62: 399–402

[29] Lilliefors, H. W. 1969. On the Kolmogorov-Smirnov test for the exponential distribution with mean unknown *Journal of the American Statistical Association* 64: 387–389

[30] Patrick Royston 1995 Remark AS R94: A remark on Algorithm AS 181: The *W* test for normality. *Applied Statistics*, 44: 547–551

[31] Patrick Royston 1982 An extension of Shapiro and Wilk's *W* test for normality to large samples. *Applied Statistics* 31: 115–124.

[32] Patrick Royston 1982 Algorithm AS 181: The *W* test for Normality. *Applied Statistics*, 31: 176–180

[33] Bartlett, M. S. (1937) Properties of sufficiency and statistical tests Proceedings of the Royal Society of London Series A 160, 268–282

[34] Pearson, Karl 1893. Contributions to the mathematical theory of evolution [abstract] *Proceedings of the Royal Society* 54 (326–330): 329–333. doi:10.1098/rspl.1893.0079. JSTOR 115538.

[35] Pearson, Karl 1895. "Contributions to the mathematical theory of evolution, II: Skew variation in homogeneous material".*Philosophical Transactions of the Royal Society* 186: 343–414. Bibcode:1895RSPTA.186..343P. doi:10.1098/rsta.1895.0010. JSTOR 90649.

[36] Pearson, Karl1901. "Mathematical contributions to the theory of evolution, X: Supplement to a memoir on skew variation".*Philosophical Transactions of the Royal Society A* 197 (287–299): 443-459 Bibcode:1901RSPTA.197..443P.doi:10.1098/rsta.1901.0023. JS-

TOR 90841.
[37] [http://en.wikipedia.org/wiki/Generalized_beta_distribution]
[38] [http://ideas.repec.org/a/ebl/ecbull/eb-11-00897.html]
[39] [http://papers.ssrn.com/sol3/papers.cfm?abstract id=2345972]
[40] http://mpra.ub.uni-muenchen.de/52344/