Valuation Biases, Error Measures, and the Conglomerate Discount

Ingolf Dittmann and Ernst Maug

ERIM REPORT SERIES *RESEARCH IN MANAGEMENT*	
ERIM Report Series reference number	ERS-2006-011-F&A
Publication	March 2006
Number of pages	48
Persistent paper URL	
Email address corresponding author	dittmann@few.eur.nl
Address	Erasmus Research Institute of Management (ERIM) RSM Erasmus University / Erasmus School of Economics Erasmus Universiteit Rotterdam P.O.Box 1738 3000 DR Rotterdam, The Netherlands Phone: + 31 10 408 1182 Fax: + 31 10 408 9640 Email: info@erim.eur.nl Internet: www.erim.eur.nl

Bibliographic data and classifications of all the ERIM reports are also available on the ERIM website:
www.erim.eur.nl

ERASMUS RESEARCH INSTITUTE OF MANAGEMENT

REPORT SERIES
RESEARCH IN MANAGEMENT

ABSTRACT AND KEYWORDS	
Abstract	We document the importance of the choice of error measure (percentage vs. logarithmic errors) for the comparison of alternative valuation procedures. We demonstrate for several multiple valuation methods (averaging with the arithmetic mean, harmonic mean, median, geometric mean) that the ranking of valuation methods is largely a function of the error measure chosen. Percentage errors give a higher weight to relative overestimates than to underestimates, and all established multiple valuation methods exhibit a positive bias according to this measure. Percentage errors lead to consequences that are not intuitive: E.g. setting company values equal to their book values often becomes the best valuation method. Logarithmic errors give equal weight to relative overestimates and underestimates and avoid unwanted consequences. With logarithmic errors, median and geometric mean are unbiased while the arithmetic mean is biased upward as much as the harmonic mean is biased downward. Measuring the diversification discount with the arithmetic mean generates a discount about twice as large as with the geometric mean or the median, whereas the harmonic mean leads to a diversification premium.
Free Keywords	Valuation, Conglomerate Discount, Financial Ratios
Availability	The ERIM Report Series is distributed through the following platforms: Academic Repository at Erasmus University (DEAR), DEAR ERIM Series Portal Social Science Research Network (SSRN), SSRN ERIM Series Webpage Research Papers in Economics (REPEC), REPEC ERIM Series Webpage
Classifications	The electronic versions of the papers in the ERIM report Series contain bibliographic metadata by the following classification systems: Library of Congress Classification, (LCC) LCC Webpage Journal of Economic Literature, (JEL), JEL Webpage ACM Computing Classification System CCS Webpage Inspec Classification scheme (ICS), ICS Webpage

Valuation Biases, Error Measures, and the Conglomerate Discount[1]

Ingolf Dittmann[2] Ernst Maug[3]

This Draft: February 14, 2006

[1]We thank Christian Weiner for extensive research assistance. We gratefully acknowledge financial support from the Rudolf von Bennigsen-Foerder Foundation and from the Deutsche Forschungsgemeinschaft through the SFB 649 "Economic Risk."

[2]Erasmus University Rotterdam, P.O. Box 1738, 3000 DR, Rotterdam, The Netherlands. Email: dittmann@few.eur.nl. Tel: +31 10 408 1283.

[3]Universität Mannheim, D-61381 Mannheim, Germany. Email: maug@cf.bwl.uni-mannheim.de, Tel: +49 621 181 1952.

Valuation Biases, Error Measures, and the Conglomerate Discount

Abstract

We document the importance of the choice of error measure (percentage vs. logarithmic errors) for the comparison of alternative valuation procedures. We demonstrate for several multiple valuation methods (averaging with the arithmetic mean, harmonic mean, median, geometric mean) that the ranking of valuation methods is largely a function of the error measure chosen. Percentage errors give a higher weight to relative overestimates than to underestimates, and all established multiple valuation methods exhibit a positive bias according to this measure. Percentage errors lead to consequences that are not intuitive: E.g. setting company values equal to their book values often becomes the best valuation method. Logarithmic errors give equal weight to relative overestimates and underestimates and avoid unwanted consequences. With logarithmic errors, median and geometric mean are unbiased while the arithmetic mean is biased upward as much as the harmonic mean is biased downward. Measuring the diversification discount with the arithmetic mean generates a discount about twice as large as with the geometric mean or the median, whereas the harmonic mean leads to a diversification premium.

JEL Classification: G10, G34

Keywords: Valuation, Conglomerate Discount, Financial Ratios

1 Introduction

This paper analyzes a methodological question that turns out to be of primary importance for valuation research: which error measure should be used when comparing alternative valuation methods. Of the 14 papers on horse races of multiples and other valuation methods (see the survey in Table A in the appendix), nine measure valuation accuracy based on the percentage difference between estimated values and market values, whereas another five use log errors, defined as the logarithm of the ratio of the estimated value to the market value. Only two of the articles that use percentage errors motivate their choice, and no paper explicitly recognizes the choice of error measure as a critical decision in the research design.[1] Also, no paper reports results for both error measures. In this paper we show that the researcher's choice of error measure is critical. This choice determines whether a valuation method produces a bias or not and therefore predisposes the conclusion in favor of certain types of valuation methods.

We demonstrate our point by investigating a question that is of independent interest. We compare four methods for averaging multiples: the arithmetic mean, median, harmonic mean, and the geometric mean. The use of averaging procedures in academic research does not reveal a consensus: median, arithmetic mean, and harmonic mean are used by different researchers, and some papers use several averaging procedures simultaneously without providing the reader with explicit guidance as to which one is preferable.[2] Several researchers have recently argued in favor of the harmonic mean as the best choice as it corrects for the apparent upward bias of the arithmetic mean.[3] This paper analyzes how the preference for one averaging method over another is predicated on the use of different error measures. In particular, we show that the harmonic mean is just as biased as the arithmetic mean if we

[1] Alford (1992) argues that absolute percentage errors put equal weight on positive and negative errors. Beatty, Riffe, and Thompson (1999) also provide an explicit justification.

[2] Beatty, Riffe, and Thompson (1999) use three error measures, reporting means and medians alongside each other is more standard. Only Herrmann and Richter (2003) use the geometric mean.

[3] Baker and Ruback (1999) argue that the harmonic mean is an ML-estimator in a model where valuation errors are normally distributed. Liu, Nissim and Thomas (2002a, 2002b) provide a derivation that supports the use of the harmonic mean as a viable and unbiased estimator. Beatty, Riffe, and Thompson (1999), Bhojraj and Lee (2002), and Herrmann and Richter (2003) also use the harmonic mean.

use log errors and that it owes its popularity only to the use of percentage errors.[4]

While we advocate the use of log errors in this paper, many researchers find percentage errors more intuitive. Error measures ultimately depend on the loss function of the researcher or analyst who needs to choose a valuation procedure. As such, error measures are subjective and not debatable. We still argue that log errors are more desirable than percentage errors by demonstrating that the usage of percentage errors produces some results that are not intuitive. In particular, we show that an *ad hoc* method that ignores all comparable information and sets the predicted market value equal to the firm's book value turns out to be just as good or even better than any of the four comparable procedures if percentage errors are used. Moreover, ignoring all information and setting the predicted firm value equal to \$1 leads to comparatively low percentage errors and – with some parameter constellations – turns out to be the best valuation method when judged by percentage errors.

Our explanation for these results is simple. Percentage errors penalize overvaluations more than undervaluations. While undervaluations larger than 100% are impossible by virtue of limited liability, overvaluations are not limited and often much more extreme than 100%. An overvaluation by a factor of 3 produces a percentage error of 200%, whereas an undervaluation by the same factor produces a percentage error of –67%, a number three times smaller in absolute value. As a consequence, judging valuation methods on the basis of percentage errors creates a preference for methods that avoid large overvaluations. Setting market values equal to book values severely undervalues companies on average as the market-to-book ratio is 1.9 for the typical company in our sample, but this *ad hoc* procedure avoids large overvaluations. The same is true for the more extreme – and in our view quite absurd – approach of setting company values equal to \$1. Effectively, this sets all percentage errors equal to –100% by fiat. However, all averaging methods produce percentage errors in excess of +100% between one fifth and one third of the time, and errors exceeding 200% or more are not uncommon. Setting company values equal to \$1 conveniently avoids percentage errors

[4] Note from Table A in the appendix that all researchers who use or favor the harmonic mean also base their error statistics on percentage errors.

of this magnitude.[5]

Logarithmic errors avoid these pitfalls. The log error of an overvaluation by a factor of 3 is exactly equal in magnitude to an undervaluation by this factor, and undervaluations of 100% produced by setting values equal to $1 appropriately create extremely large negative error measures. Statistically, logarithmic error distributions are more symmetric and closer to satisfying the normality assumptions often made for statistical inference.[6]

We base our analysis on a sample of 52,112 firm-year observations from 1994 to 2003 for U.S. companies and show that the distributions of standard financial ratios (market-to-book, value-to-sales, price-earnings) are heavily skewed and much better modelled by a lognormal distribution than by a normal distribution, even though both distributions are rejected in formal tests. We then develop a two-pronged approach. We first develop an analytic argument based on the lognormal distribution. We establish those results that cannot be derived analytically through Monte Carlo simulations calibrated to the moments of the empirical distribution of market-to-book ratios. We show that for log errors the arithmetic mean has a positive bias and the harmonic mean has a negative bias, and that both biases are equal in absolute value. Only the geometric mean is unbiased in all cases, which is not surprising once we recognize that the geometric mean can be obtained by first calculating the arithmetic mean of the logarithms of the relevant financial ratios and then transforming back from logarithms: this transformation is precisely what is required to neutralize the skewness of the lognormal distribution and it is somewhat surprising that the geometric mean is the only averaging procedure that did not find many followers so far.[7]

Our second approach is empirical. As empirical distributions of multiples are not exactly lognormal, we compare the two error measures using our sample of historical data. We

[5] Some researchers report median percentage errors or other percentiles of the error distribution. We comment on this practice in more detail in Sections 3 and 4 below.

[6] To the best of our knowledge, only Baker and Ruback (1999) have explicitly tested if the normality assumption applies to their sample and they could not reject it. However, note from Table A that their sample of 225 observations is rather small by the standards of the valuation literature. Kaplan and Ruback (1995), Lie and Lie (2002), and Hermann and Richter (2003) explicitly motivate the use of log errors with the skewness of percentage errors or the distributions of the underlying fundamental variables.

[7] An exception is the little known study by Herrmann and Richter (2003) who do not call it the geometric mean but the "retransformed logarithmic mean," appealing to the same argument as presented in the text.

find that the qualitative conclusions from the theory and the simulations continue to hold. In particular, the geometric mean and the median are unbiased with log errors, while all methods have a positive bias and the harmonic mean is least biased with percentage errors.

Finally, we apply our insights to the measurement of the diversification discount. The literature on the diversification discount relies exclusively on multipliers (asset and sales multipliers, and Tobin's q) to measure the fundamental value of the segments of diversified firms. As such, this literature constitutes the academic showpiece for the use of multiples and a good testing ground for multiples research.[8] We show that large estimates of the diversification discount are obtained simply from biased benchmarks generated by averaging multiples with the arithmetic mean. The median or geometric mean give more consistent and much smaller estimates of about 20%. In contrast, the harmonic mean leads to a small diversification premium of up to 3%.

The following Section 2 discusses our sample and analyzes different distributional models for multiples. Section 3 presents our conceptual analysis based on analytic proofs and simulations. Section 4 conducts the empirical analysis based on tests similar to those in Section 3. Section 5 applies our insights to the measurement of the diversification discount and Section 6 concludes. Technical material is deferred to the appendix.

2 Data and the distribution of multiples

In this section, we present our dataset and demonstrate that the frequency distribution of conventional financial ratios resembles a lognormal distribution much more than a normal distribution. We will use the assumption that financial ratios are lognormally distributed in the next section when we derive our theoretical results.

Our analysis is based on annual data from Compustat between 1994 and 2003. We select all companies domiciled in the United States whose sales and total assets both exceed \$1 million. We also require that the market value of equity four months after the fiscal year end

[8] The diversification discount literature and the valuation literature have largely proceeded independently of each other. The literature on the diversification discount begins with Lang and Stulz (1994) and Berger and Ofek (1995). We discuss this literature in further detail in Section 5 below.

is available. The four months lag ensures that the company's financial statements have been publicly available to investors and are therefore reflected in the market value. We exclude those companies where the SIC code is either not available or equals 9999 (not classifiable).[9] Finally, we delete all firms in industries (as defined by the two-digit SIC code) with less than 5 firms. We are left with a final dataset with 52,112 firm-year observations.

We focus on three multipliers:

- market-to-book ratio, defined as the market value of equity divided by the book value of equity.
- value-to-sales ratio, defined as the ratio of enterprise value to sales, where enterprise value is the market value of equity plus total debt.
- price-earnings ratio, defined as market value of equity divided by net income.

A multiple that is negative according to these definitions is set to a missing value. We also set the market-to-book ratio equal to a missing value if shareholders' equity is smaller than $1 million. We can compute the market-to-book ratio for 47,614 firm-year observations, the value-to-sales ratio for 51,899 observations, and the price-earnings ratio for 33,753 observations. Finally, we winsorize the data separately for each multiple and each year at 2.5% and 97.5%. We report descriptive statistics for all three ratios and their natural logarithms in Table 1.

[Insert Table 1 about here]

The table shows that the median market-to-book ratio in our sample is 1.87. The median value-to-sales ratio is 1.63, and the median price-earnings ratio is 17.1.[10] As usual, all

[9] We use the historical SIC code (SICH) when available. If the historical SIC code is not recorded in Compustat, we use the current SIC code (SIC).

[10] Bhojraj and Lee (2002) report a similar median market-to-book ratio of 1.84. Their mean (2.26) is substantially lower than our mean, because they delete extreme values while we winsorize them; also their sample ranges from 1982 to 1998 and excludes the high valuation years 1999 and 2000 included in our sample. Of the other comparable studies, Beatty, Riffe, and Thompson (1999) and Liu, Nissim, and Thomas (2002a) report all variables scaled by price, which is the inverse of our ratio and has different statistical properties. Alford (1992) and Cheng and McNamara (2000) work with much older samples. Lie and Lie (2002) work with different definitions (enterprise value to total assets instead of market-to-book) and only with 1998 data.

distributions are highly skewed and means substantially exceed medians. Table 1 also reports considerable positive excess kurtosis for all three distributions, i.e. all distributions have fatter tails than the normal distribution. In contrast, the distributions of the logarithms of all ratios are more symmetric and show less excess kurtosis than the original distributions, although some skewness still remains.

Figure 1 graphs the standardized distributions of the three multiples and their log transformations. We standardize each observation by deducting the industry-year mean and dividing by the industry-year standard deviation. Hence, we allow for different means and variances across industries and years. In the left column of Figure 1, we always compare the empirical distribution of the untransformed data (solid line) with the best fits obtained for the normal distribution (dotted line). In the right column we compare the distribution of the logarithmic transformations of the original data with the normal distribution (dotted line).

Clearly, the lognormal distribution appears to be a better model than the normal distribution for all three financial ratios. From a visual inspection of the graph, the lognormal distribution appears to be a reasonable model for the market-to-book ratio and the value-to-sales ratio, but a somewhat less convincing model for the price-earnings ratio. However, even for the market-to-book ratio and the value-to-sales ratio, the lognormal distribution does not describe the empirical distribution perfectly. For all ratios, the empirical distributions are more skewed and exhibit fatter tails than the lognormal distribution. We test the fit of the distributions to the data more formally by applying three standard tests for normality to the ratios and their log transformations.[11] These are reported in Table 2.

[Insert Table 2 about here]

All three tests reject the normal distribution as the correct model for the distributions

[11] All three tests compare the empirical distribution function with the normal distribution function where the mean and the variance are estimated from the sample. The Kolmogorov-Smirnov test is based on the maximum absolute distance between the two distributions. In contrast, the Cramer-von Mises test and the Anderson-Darling test are based on the expected sum of squared distances under the normal distribution function. While the Cramer-von Mises test gives equal weight to all observations, the Anderson-Darling test gives higher weight to the tails of the distribution. See D'Agnostino and Stephens (1986), p.100.

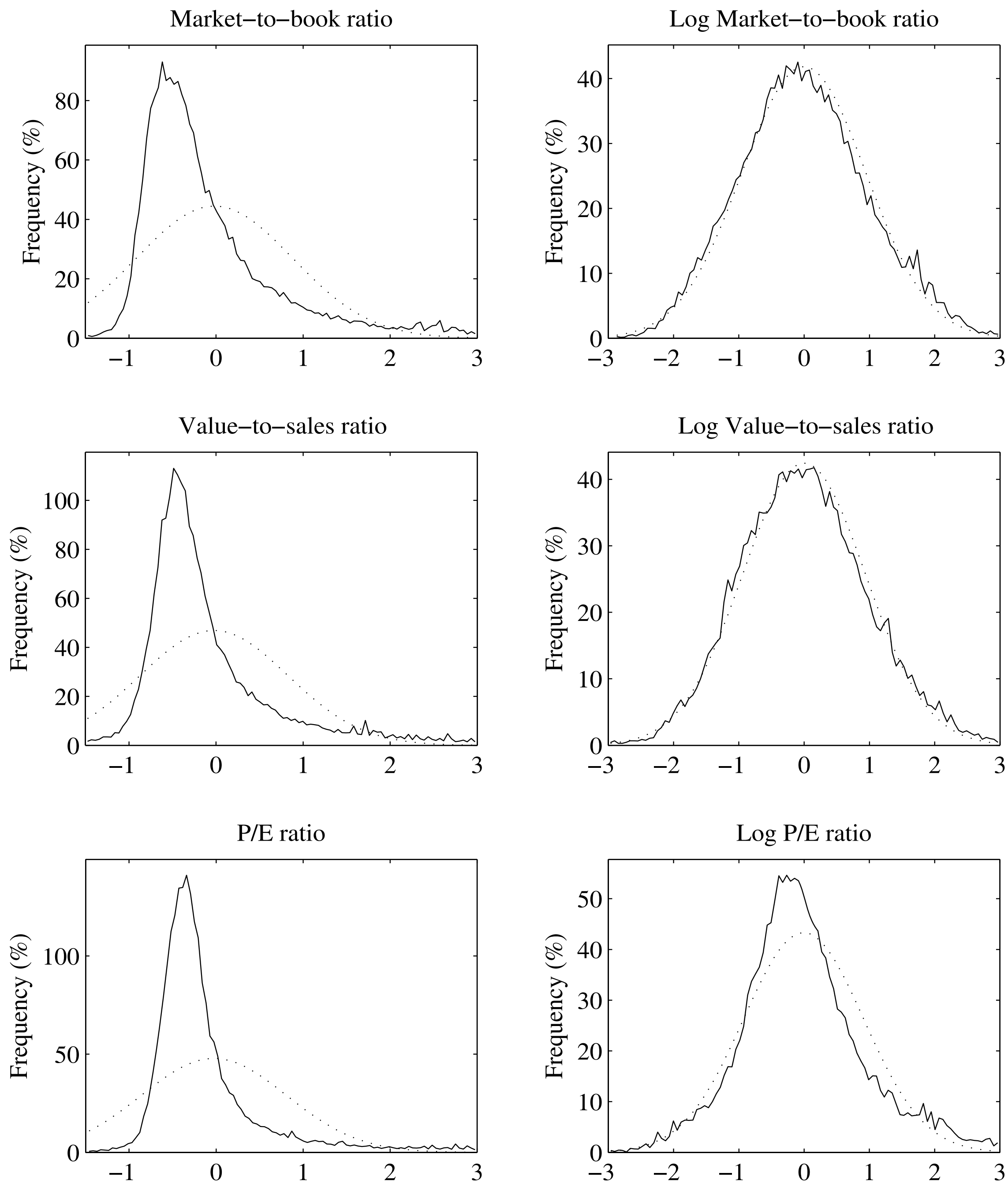

Figure 1: The figure shows the empirical distributions of the financial ratios (left column) and their logarithmic transformations (right column), calculated with 100 histogram intervals. In each plot, the solid line shows the actual distribution of the data and the dotted line shows the density function of a normal distribution with mean and variance fitted to the data. All observations have been standardized by deducting the industry-year mean and then dividing by the industry-year standard deviation, where the industry is given by the two-digit SIC code. The data have been truncated to generate meaningful plots.

of all three ratios as well as for their logarithmic transformations. Hence, we also reject the lognormal distribution as the correct distributional model. This is unsurprising given that we have a very large dataset, so that the tests have high power. However, a closer look at the test statistics supports the same conclusion already suggested by Figure 1: the lognormal distribution is a much better model for all three financial ratios than the normal distribution, given that the data are highly skewed. The Kolmogorov-Smirnov test statistic falls by between 70% (price-earnings) and 90% (value-to-sales), the Cramér-von Mises and Anderson Darling test statistics even fall by up to 99% for the value-to-sales ratio. In all cases the relative improvement is largest for the value-to-sales ratio and smallest for the price-earnings ratio, which is consistent with our visual inspection of these distributions in Figure 1. We therefore conclude that the lognormal distribution works best for the value-to-sales ratio and worst for the price-earnings ratio.[12]

There is also a theoretical reason why some financial ratios are better approximated by the lognormal distribution than others. Variables like market value, book value, sales, total assets, or the number of employees can only be positive and are measures of firm size. If these variables are lognormally distributed (which is a common assumption in statistical applications), then the ratio of two of these variables is also lognormally distributed. This argument does not hold for performance measures, like net income or EBIT. These variables can become zero or negative, and it is not clear whether the truncated distribution obtained by discarding negative values can be approximated well by a lognormal distribution.

In the next section, we will work with the assumption that financial ratios are lognormally distributed. This assumption allows us to generate strong theoretical results and to perform easily parameterized simulations. The evidence presented so far justifies this assumption but also makes clear that it constitutes a rather strong simplification of reality. We will therefore check all our theoretical results with the true empirical distribution in Section 4. Alternatively, we could search for a better statistical model than the lognormal distribution and base our conceptual analysis on such a model. We do not pursue this and argue below

[12] We also repeated these tests for the book-to-market ratio (i.e. the inverse of the market-to-book ratio) and obtained similar results which we therefore do not report in the tables.

that this would probably have little impact on our results.

3 Conceptual Analysis

In this chapter we use theoretical arguments and Monte Carlo simulations to investigate the relationship between error measures (percentage error or logarithmic error) on the one hand and multiple valuation methods (median, arithmetic mean, harmonic mean, geometric mean) on the other. In particular, we demonstrate that the optimal choice of averaging method crucially depends on the error measure chosen by the researcher.

3.1 Theoretical Analysis

We wish to value some company j on the basis of some financial ratio denoted by x. We use subscripts to refer to individual companies, so x_j is the value of x for company j. We will refer to company j also as the target firm. The multiple x is the ratio of the market value of the company, denoted by MV, to some base B, where the base may be the book value of the assets, sales revenues, or a measure of firms' profitability like EBIT, EBITDA, or earnings, so that $x_j = MV_j/B_j$. The literature has used different definitions of market value and a variety of different bases that do not concern us in this section beyond their statistical properties.

The financial ratio x can be measured for a set of companies that are deemed comparable to company j by the researcher or analyst. We denote this set by $I_j = \{1, ..., n\}$, where $j \notin I_j$. We assume throughout this subsection that the $\{x_i | i \in I_j\}$ and x_j have been drawn independently from an identical distribution.[13] Then we compute an average financial ratio $\bar{x}_j$ across all comparable firms $i \in I_j$ and multiply it by firm j's base B_j in order to obtain an estimate of firm j's market value:

$$\widehat{MV}_j := \bar{x}_j \times B_j \ . \tag{1}$$

[13] In addition, we assume that the set I_j contains at least 2 elements that differ from one another. We maintain the independence assumption only for expositional convenience. All our results can also be derived under weaker assumptions that allow for dependence between the x_j.

We consider four different averaging methods:

$$\text{Arithmetic Mean:}\quad \bar{x}_j^A = \frac{1}{n}\sum_{i=1}^{n} x_i \tag{2}$$

$$\text{Harmonic Mean:}\quad \bar{x}_j^H = \frac{1}{\frac{1}{n}\sum_{i=1}^{n}\frac{1}{x_i}} \tag{3}$$

$$\text{Geometric Mean:}\quad \bar{x}_j^G = \prod_{i=1}^{n} x_i^{1/n} = \exp\left\{\frac{1}{n}\sum_{i=1}^{n}\ln(x_i)\right\} \tag{4}$$

$$\text{Median:}\quad \bar{x}_j^M = Median\,(x_i, i \in I_j) \tag{5}$$

The second expression in (4) shows that the geometric mean can be interpreted as a retransformed arithmetic mean of the logs of the multiples x_i.

Proposition 1 collects some basic results about the relative size of these four averages. We give a short proof in the appendix but do not credit ourselves with first establishing these results. They follow directly from Jensen's inequality and the law of large numbers.

Proposition 1 ***(Means):*** *(i) The arithmetic mean always results in a higher market value estimate than the geometric mean or the harmonic mean, and the harmonic mean always results in a lower market value estimate than the geometric mean:*

$$\bar{x}_j^H < \bar{x}_j^G < \bar{x}_j^A.$$

The relative position of the median, $\bar{x}_j^M$, depends on the distribution of the ratios.

(ii) If the ratios have been drawn from a symmetric distribution, the median is close to the arithmetic mean in sufficiently large samples. More formally: $\left(\bar{x}_j^M - \bar{x}_j^A\right) \to 0$ as $n \to \infty$.

(iii) If the ratios have been drawn from a lognormal distribution, the median is close to the geometric mean in sufficiently large samples. More formally: $\left(\bar{x}_j^M - \bar{x}_j^G\right) \to 0$ as $n \to \infty$.

Hence, an analyst who uses the harmonic mean will *always* arrive at lower value estimates than her colleague who works with the geometric mean, given that they use the same set of comparable companies. Both will obtain lower estimates than a third analyst using the arithmetic mean. The differences between these three measures and the median are larger if the variation in the sample is large. So for a set of comparables with little variation, the four methods arrive at similar results. Note also that part (i) of Proposition 1 makes no assumption on the distributions of the x_j, and we need the assumption of a lognormal distribution only for part (iii).

In order to compare the precision of the four averaging methods, we need an error measure that defines what we mean by 'relative precision.' Two error measures are commonly used in the literature on company valuation:

$$\text{Percentage Error:} \quad e_{pct}(j) = \frac{\widehat{MV}_j - MV_j}{MV_j} = \frac{\bar{x}_j \times B_j \ - x_j \times B_j}{x_j \times B_j} = \frac{\bar{x}_j}{x_j} - 1 \ , \tag{6}$$

$$\text{Log Error:} \quad e_{\log}(j) = \ln\left(\frac{\widehat{MV}_j}{MV_j}\right) = \ln\left(1 + e_{pct}(j)\right) = \ln\left(\frac{\bar{x}_j}{x_j}\right) \ . \tag{7}$$

$>$From now on we will suppress the reference to firm j for errors for ease of exposition. Note that the second transformation in (6) and (7) shows that the errors and their statistical properties are independent of the multiple base B_j of the target firm.

In order to select optimal valuation procedures we investigate statistics of these error measures. Specifically, we are interested in the bias $E(e_j)$ and the mean squared error $E\left(e_j^2\right)$ or the Root Mean Squared Error (RMSE). Some of the literature following Alford (1992) has also looked at the mean or median absolute error and we will therefore also look at $E(|e_{pct}|)$. While it is true that the two error functions do not differ significantly for small errors as we have $e_{pct} = e_{\log} + O\left((e_{\log})^2\right)$ from a first order Taylor expansion, valuation errors with multiples are often large and then the two error functions generate some surprising differences. Proposition 2 summarizes our theoretical results about the biases of the different valuation methods with the two error measures:

Proposition 2 ***(Errors):*** *Assume that all comparable ratios x_i and x_j are distributed log-normal with parameters μ and $\sigma^2 > 0$.*

(i) For ***percentage errors,*** *the geometric mean and the arithmetic mean are both biased*

upward. The bias of the arithmetic mean is stronger than that of the geometric mean.

$$0 < E\left(e^G_{pct}\right) < E\left(e^A_{pct}\right) \ .$$

The expected error of the harmonic mean is smaller than that of the geometric mean: $E\left(e^H_{pct}\right) < E\left(e^G_{pct}\right)$*, and the median is also biased upward in large samples.*

(ii) For ***log errors,*** *the geometric mean is unbiased, the arithmetic mean is biased upward, and the harmonic mean is biased downward.*

$$E\left(e^H_{\log}\right) < E\left(e^G_{\log}\right) = 0 < E\left(e^A_{\log}\right) \ .$$

In absolute terms, the harmonic and the arithmetic mean are equally biased: $E\left(e^H_{\log}\right) = -E\left(e^A_{\log}\right)$*. The median is unbiased in large samples.*

The arithmetic mean has a positive bias for both definitions of valuation errors. The distribution of x_i is skewed and the arithmetic mean gives equal weight to all observations, including large outliers that necessarily occur with skewed distributions. The geometric mean is unbiased for log errors for reasons that are intuitive from looking at the second transformation in equation (4). The geometric mean is the retransformed arithmetic mean, applied to the logarithms of the financial ratios x_i, and $\ln(x_i)$ is distributed normal. Hence, this implicit logarithmic transformation removes the skewness of the original distribution and the transformed distribution is symmetric, which is exactly what is required for generating unbiased logarithmic errors. However, unbiased logarithmic errors are not equivalent to unbiased percentage errors. From (7) observe that $e_{pct} = \exp(e_{\log}) - 1$ and this convex transformation gives positive errors a larger weight than negative errors, creating an upward bias. By comparison the harmonic mean is biased downward for logarithmic errors. The skewness of the lognormal distribution implies that large outliers are balanced by a larger number of very small observations, which then create very large numbers $1/x_i$ from the definition of the harmonic mean. This inflates the denominator of $\bar{x}^H_j$ and biases the valuation errors downward. Interestingly, the harmonic mean is biased downward just as much as the arithmetic mean is biased upward.

There are two questions that we cannot settle by way of analytic proofs. Firstly, we cannot infer if $E\left(e^H_{pct}\right)$ is positive or negative, or, indeed, if $E\left(e^H_{pct}\right)$ is larger or smaller in absolute value than $E\left(e^G_{pct}\right)$ and $E\left(e^A_{pct}\right)$. Secondly, we cannot derive the properties of the median in small samples or determine analytically what a large sample is. We will address these questions by way of a Monte Carlo simulation. The simulation also allows us to quantify how severe the different biases are.

3.2 Monte Carlo simulations

Simulation methodology. In our theoretical analysis we only describe the four averaging methods (median, arithmetic mean, harmonic mean, and geometric mean) in terms of their biases. A good valuation method should have no or only a small bias, but it should also have a low dispersion. In this section, we compare the four different averaging methods and the two error measures using simulated data. The Monte Carlo approach allows us to obtain the full error distribution and to calculate more statistics than we can derive theoretically. On the other hand, simulations allow us to abstract from many complicated features of observed data, that are caused e.g. by sampling errors or accounting conventions.

In addition to the four averaging methods, we also consider two benchmark valuation procedures. These procedures make no use of comparable information and we would expect that any valuation procedure that incorporates more information should also generate lower errors. In particular:

1. We set the value of the target company arbitrarily equal to its book value, which amounts to setting the market-to-book value of the target company equal to one. We therefore call this strategy in the tables below "MTB = 1." This is clearly a very rough and imprecise valuation method that is based on only one piece of accounting information. Note that this method is biased downward as the median and mean market-to-book ratio are substantially larger than one in our sample. We would not wish to use a procedure that does as poorly or possibly worse than this.

2. We set the value of the company arbitrarily equal to a very small value close to zero. For this we choose \$1. We refer to this procedure as "Value = \$1" in the tables. This procedure is even worse than the first as it relies on no company information at all and any reasonable procedure should find it easy to beat this benchmark.

Both dummy-procedures lead on average to undervaluations and large biases. However, they generate very little dispersion and avoid overvaluations. The Value=\$1 procedure also avoids errors of more than 100% in absolute value.

In practice, analysts use a set of comparable firms - rather than all available firms - in order to predict the market value of another, similar firm. Following much of the literature, we choose the same industry as determined by the 2-digit SIC code as the appropriate peer group.[14] In our sample, we have 608 industry-years with 5 or more firms each, i.e. on average 61 industries per year. The average industry contains 78 firms, and the median industry 38 firms. We therefore simulate industries with 40 firms. Later we will also consider industries with 5 or 200 firms.

The basic case. We perform 100,000 runs. For each run, we draw 40 market-to-book ratios from a lognormal distribution with $\mu = 0.597$ and $\sigma^2 = 0.801$. These parameters are the average industry mean and the average within-industry standard deviation of the log market-to-book ratio across the 608 industries in our sample. The standard deviation is somewhat lower than the overall standard deviation in our sample (0.882, see Table 1), so choosing firms from the same industry reduces the dispersion as it removes between-industry variation. The mean is also different because we average first within each industry and then across industries, which gives firms in smaller industries a larger weight. Likewise, we draw 40 book-values from a lognormal distribution with $\mu = 18.312$ and $\sigma^2 = 1.828$. Again these parameters have been estimated from our dataset. The firm's market value is then given by the product of market-to-book ratio and book value. In the next step, we estimate the market value of each of the 40 firms with each of the six methods considered. For the four

[14] See Alford (1992), Cheng and McNamara (2000), and Bhojraj and Lee (2002) for an analysis of more sophisticated methods to choose comparables.

averaging methods, we use (2) - (5). For MTB = 1 or Value = \$1, we set the estimated firm value equal to the book value or equal to \$1, respectively.[15] Finally, we calculate the resulting percentage errors and the resulting log errors. For each valuation procedure and each error type we thereby obtain $40 \times 100,000 = 4,000,000$ errors. Table 3 reports some statistics for the error distributions, Panel A for log errors, Panel B for percentage errors, and Panel C for absolute percentage errors.

[Insert Table 3 about here]

The table corroborates our theoretical results from Proposition 1: The bias for percentage errors is highest for the arithmetic mean (0.899), is lowest and almost zero for the harmonic mean (0.023), with the geometric mean and the median about halfway in between (0.389 and 0.395, respectively). We therefore reproduce the result stated in the previous literature that the harmonic mean dominates all other valuation methods in terms of percentage errors. In terms of log errors, the geometric mean and the median are unbiased, and, with a bias of 0.308, the arithmetic mean is biased upward just as much as the harmonic mean is biased downward (see Proposition 2(ii)). Interestingly, while the pattern of *mean* errors (i.e. bias) differs significantly between percentage errors and log errors, the *median* errors generate the same message for percentage errors and log errors: they are zero for the median and the geometric mean, positive for the arithmetic mean, and negative for the harmonic mean.

However, we are not only interested in the bias of valuation procedures but also in the dispersion of valuation errors. We measure these by the standard deviation and the root mean squared error (RMSE). The RMSE combines bias and standard deviation in a convenient way and can be rationalized from minimizing a quadratic loss function. The squared RMSE is equal to the variance plus the squared bias. For many applications in practice (e.g., assessing the purchase prices of unlisted companies, IPO pricing), the bias may be more important than is reflected in RMSEs. For example, for a successful acquisition strategy it may be more important to avoid consistently overpaying for acquisition targets. We therefore also report

[15] The book value is relevant only for the "Value=\$1" method. For all other methods, the book value appears in the numerator and the denominator of the two error measures and therefore cancels. See (6) and (7) and note that for MTB=1 we have $\overline{x}_j = 1$.

biases and standard deviations throughout, as different applications may warrant different weights for the bias and dispersion of valuation methods.

For percentage errors the RMSE generates a remarkable result: Estimating the target firm's market value by its book value (MTB=1) clearly outperforms all other valuation methods. Even the more extreme *ad hoc* procedure of setting the target firm's market value to \$1 (Value=\$1) turns out to be practically as good as using the harmonic mean. The reason is that percentage errors are bounded from below at -100% but not bounded from above. If errors on the unlimited upside are severe, methods that undervalue firms on average (or even set the error equal to the lower bound as Value=\$1 does) appear to be preferable. The last column of Table 3 shows that, for a third of the firms, the percentage error of the arithmetic mean exceeds 100%, while the median and the geometric mean lead to percentage errors in excess of 100% for one fifth of all firms. These high overvaluations are largely avoided by MTB=1 and completely eliminated by Value=\$1.

Log errors, on the other hand, have an unlimited downside and penalize undervaluations as much as overvaluations. Table 3A shows that the standard deviations of these log errors are virtually identical for all four averaging procedures, so that the differences in the RMSEs are generated entirely by the differences in bias. The Value=\$1 procedure is heavily and appropriately penalized for the extreme undervaluations it generates. The MTB=1 procedure generates the lowest dispersion but still has a higher RMSE than all comparables-based procedures because of its large downward bias, which is assessed at 60% in terms of log errors, but only 24% in terms of percentage errors.

The fact that over- and undervaluations are treated asymmetrically by percentage errors is also reflected in the skewness of valuation errors. Errors obtained with percentage errors are highly skewed while errors obtained with log errors are virtually symmetric. We suspect that the skewness of percentage errors caused by the limited downside and unlimited upside is the reason why many researchers who work with percentage errors or absolute percentage errors report medians (and sometimes other percentiles) rather than means of the error distribution. This approach ignores the large incidence of extreme overvaluations, however.

Absolute errors were used by some researchers (e.g. Alford, 1992) and behave broadly similarly to percentage errors, but are somewhat more extreme as Panel C of Table 3 shows. They are more skewed, and means and medians are both larger compared to percentage errors, whereas the RMSE and the proportion of errors exceeding 100% is identical by construction. Statistics of absolute errors therefore generate very similar results to those of percentage errors, and we therefore do not report results for simulated absolute percentage errors in the following tables.

Simulations for small industries. So far we have analyzed the relative performance of the four averaging methods when applied to a group of 40 comparable companies, because the median industry size in our sample is 38. We now want to check whether our results are robust with respect to the number of comparables used. Table 4 reports simulation results for small industries with only 5 firms, so that each firm is valued by 4 comparable companies. This accords with the common practice of requiring a minimum number of 4-5 comparable firms in an industry. Even some of the two-digit SIC industries in our sample are small (10% of the 608 industry-years in our sample contain between 5 and 10 firms), so analyzing the impact of industry size is important.

[Insert Table 4 about here]

The results of Table 4 are broadly similar to those of Table 3 and all remarks about rankings in terms of bias and RMSE still hold with two notable exceptions. Firstly, for percentage errors we now report a significant overvaluation of 21% even for the harmonic mean. As a result, the harmonic mean generates almost twice the RMSE as MTB=1, and Value=$1 now clearly dominates all averaging methods according to the RMSE. The reason is that all valuations become less precise: all standard deviations are markedly higher in Table 4 than in Table 3. Also the frequency of outliers is generally higher in smaller industries. Secondly, the median is not unbiased anymore for log errors and is now biased upward by 4.6% (for 100,000 runs this is statistically significant at all conventional levels). We observe that Proposition 1 only states that the median is unbiased in large samples, so our simulations suggest that 40 firms is large enough for the law of large numbers to apply, while 5 firms is not. We investigate this further by plotting the bias of the geometric mean and the median against industry size in Figure 2.

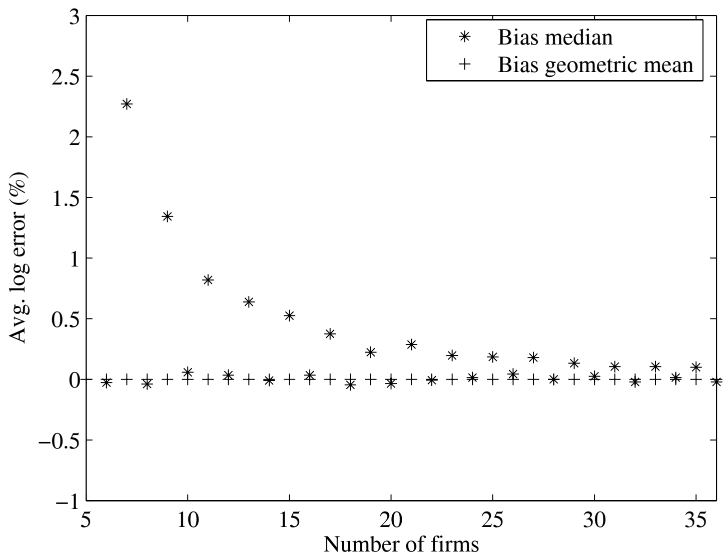

Figure 2: The figure shows simulation results for the bias of the median and the geometric mean in percent as a function of industry size (in terms of number of firms). The simulation parameters other than the number of firms are those stated in the text.

Figure 2 shows that the bias of the median is a feature of small industries and that it declines to biases below 1% for industries with more than 10 firms. Observe from the figure that the median is unbiased for odd numbers of comparables, i.e. for even numbers of firms in an industry, as the industry also contains the target firm. If the number of comparables is even, however, the median is determined by *arithmetic* averaging of the two central observations, so that the bias of the arithmetic mean carries over to the computation of the median. In the extreme case of only two comparables, arithmetic mean and median are identical. For moderately large industries these effects are not important anymore and the median is as unbiased as the geometric mean for log errors (respectively, median percentage errors).[16]

[16] The median would be approximately unbiased also for small even numbers of comparables, if the two central observations would be averaged with the geometric mean instead of the arithmetic mean. This is non-standard however.

We also performed similar simulations for large industries with 200 firms, because approximately 10% of our 608 industry-years contain 200 firms or more. It turns out that the impact of larger industry size is, somewhat predictably, the opposite of smaller industry size and differs only little from our base case with 40 firms reported in Table 3. We therefore do not report these results.

[Insert Table 5 about here]

Simulations for heterogeneous industries. A key variable in our analysis is the dispersion of the market-to-book ratio within industries. A good valuation procedure should be able to cope with industries that exhibit large dispersions of market-to-book ratios. We therefore also perform a comparative static analysis by increasing the standard deviation of the industry market-to-book ratio by 50% from 0.80 to 1.20. We choose a rather extreme increase in the standard deviation (only 2% of the 608 industry-years in our sample have a standard deviation that exceeds 1.2) in order to better demonstrate the effects of increased dispersion. Table 5 reports the results. In many ways, an increase in standard deviation magnifies the effects we have discussed previously. The bias, the standard deviation and the proportion of errors exceeding 100% increase dramatically for all valuation methods and all definitions of valuation errors. A notable exception is the geometric mean and the median with log errors which remain unbiased. Interestingly, even the MTB=1 procedure now overvalues the target firm with percentage errors (on average by 11.5%), confirming our result above that overvaluation is a feature of the percentage error and not informative about the valuation procedures themselves. Note that MTB=1 still heavily underestimates company values according to log errors. Moreover, with percentage errors Value=$1 now dominates the arithmetic mean, the median, and the geometric mean in terms of bias and RMSE.

We can now summarize our results from the theoretical analysis as follows. Percentage errors may seem more intuitive than log errors at first glance, but they lead to a number of counterintuitive results. The reason is that percentage errors penalize overvaluations more than undervaluations because the percentage error of undervaluations is limited by

-100% whereas errors for overvaluations are unlimited. As a consequence, all reasonable valuation procedures - including the harmonic mean and MTB=1 - have a positive bias with percentage errors if the variance of the relevant financial ratio is sufficiently high. Therefore, percentage errors favor valuation methods that consistently lead to low market values: Even setting the value of the firm to $1 appears to be a sensible valuation method if dispersion is high. By comparison, log errors are symmetric with respect to equal relative deviations and therefore avoid these distortions. Based on log errors, averaging multiples using harmonic means biases valuations downward, averaging with arithmetic means biases them upward, whereas medians and geometric means lead to unbiased valuations. There is little to choose between medians and geometric means, except for small industries with few firms, where the geometric mean dominates. Altogether, we interpret this as a strong argument against the widespread usage of percentage errors to compare valuation methods, as this seems to tilt the playing field in favor of valuation methods that tend to undervalue companies.

4 Empirical analysis

The conceptual argument of the previous section is based on the model of a lognormal distribution. The analysis in Section 2 above shows that the lognormal distribution is superior to the normal distribution, but is still rejected by the statistical tests reported in Table 2. We therefore compare the six valuation methods (four averaging methods and two *ad hoc* methods) on our dataset (see Section 2). This allows us to check whether our simulation results continue to hold for real world data.

For all firms in the dataset we select a set of at least five comparables from the same industry. Following Alford (1992), we use a slightly more complicated, but also more realistic method to select comparables based on SIC levels than we did in the Monte Carlo simulations: We start at the 4-digit SIC level. If we cannot find at least five comparable firms, we proceed to the 3-digit SIC level and, likewise, to the 2-digit and 1-digit SIC level, where we can match all remaining firms. We repeat this for every year from 1994 to 2003 and compute errors and statistics for the pooled sample as before. Table 6 reports the results for the market-to-book ratio.[17]

[Insert Table 6 about here]

Table 6 largely corroborates our results from the previous section. The arithmetic mean generates large positive biases for all definitions of valuation errors; the bias in terms of percentage errors is positive for all methods based on comparables; and the harmonic mean creates a downward bias when evaluated on the basis of either mean or median log errors or median percentage errors. Consistent with the previous literature, the harmonic mean is favored by mean percentage errors. For log errors, the geometric mean is unbiased and the median exhibits a small but statistically highly significant negative bias. Finally, the upward bias of the arithmetic mean is almost equal to the downward bias for the harmonic mean as predicted by Proposition 1.

Figure 3 contains eight graphs that show the error distributions: each of the four rows corresponds to one of the four averaging methods (arithmetic mean, median, geometric mean, and harmonic mean). The left graphs give the distributions of percentage errors, whereas the right graphs show the distributions of log errors. Clearly, all distributions of percentage errors are highly skewed. Also, all distributions exhibit a significant proportion of percentage errors that exceed 100%. The graphs confirm our intuition that errors based on log transformations are much closer to the model of a normal distribution than percentage errors. Apparently, the log transformation is successful in generating a symmetric distribution centered around zero for the median and the geometric mean, whereas the distributions of log errors for the arithmetic mean and the harmonic mean are not symmetric. The distribution of log errors for the arithmetic mean has an extremely fat right tail, and its mode is clearly positive. The harmonic mean, on the other hand, exhibits a fat left tail and a negative mode. Hence,

[17] In Tables 6, 7, and 8, we exclude observations with valuation errors larger than 1000 (i.e. 100,000%) under percentage errors before calculating the statistics shown in the Table. Accordingly we exclude one observation for Table 6 (market-to-book), one observation for Table 7 (enterprise-value-to-sales), and 13 observations for Table 8 (price-earnings). These obvious outliers heavily influence standard deviations and RMSEs under percentage errors. On the other hand, they have only little effect on the numbers reported under log errors.

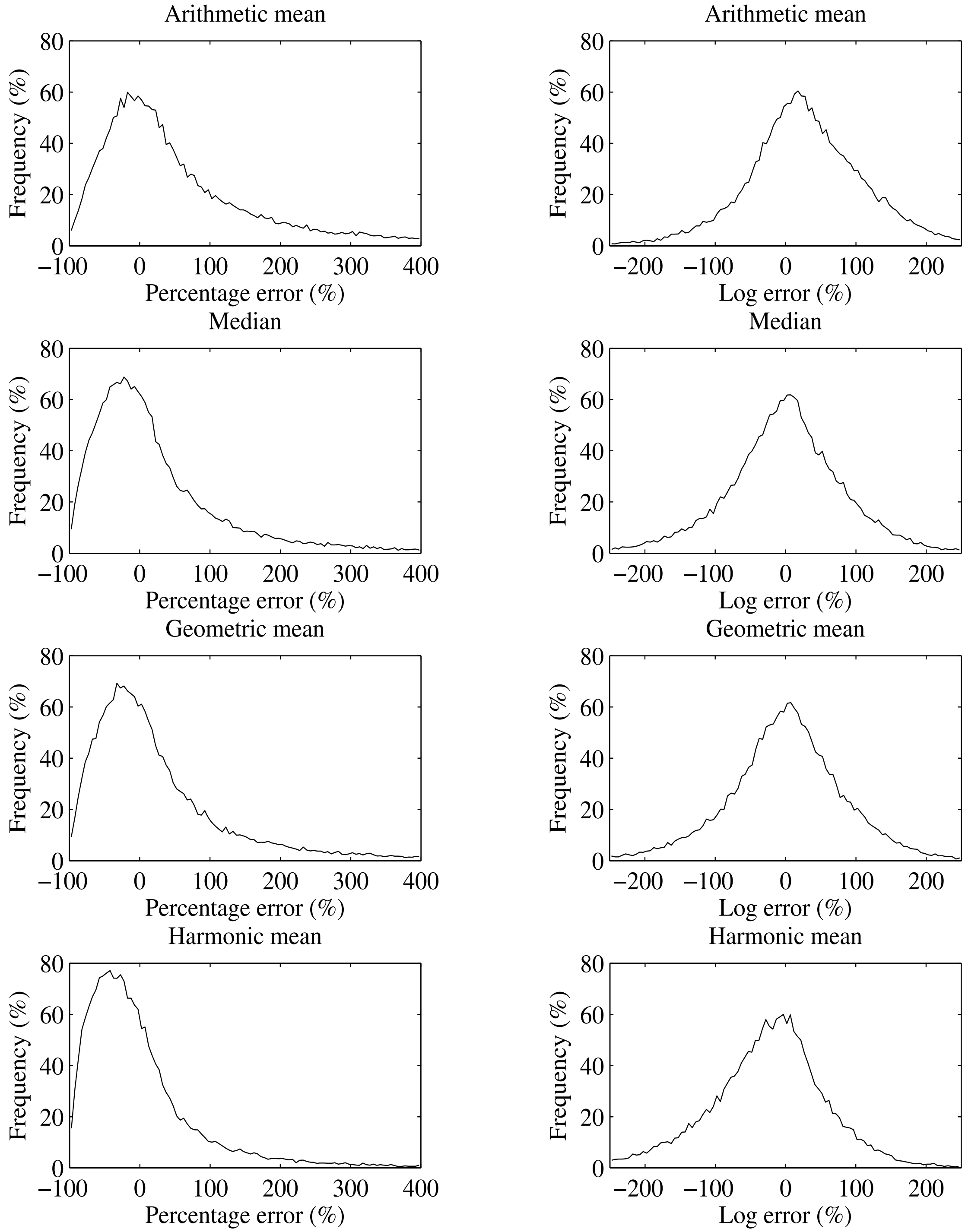

Figure 3: The figure shows the empirical error distributions based on the market-to-book ratio, calculated with 100 histogram intervals from 47,614 firm-year observations from 1994-2003. The left column shows the percentage errors, whereas the right column shows log errors. The rows correspond to one averaging method each. The data have been truncated to generate meaningful plots.

the harmonic mean generates more undervaluations, whereas the arithmetic mean generates more overvaluations.

The dispersion measures (RMSE and standard deviation) are much larger for the empirical distribution than suggested by the simulations of the previous section, evidently a consequence of the much fatter tails of the empirical distribution of market-to-book ratios. In terms of the RMSEs and standard deviations of percentage errors, we also obtain the same ranking of valuation methods as before, with the *ad hoc* methods dominating all comparable methods and the harmonic mean as the best comparable method. However as discussed on page 16, we are sceptical that the RMSE provides a useful weighting of dispersion and bias. Still, even in terms of bias the MTB=1 method dominates the median, geometric mean, and the arithmetic mean in terms of percentage errors (Panel B) and the bias for MTB=1 is very similar in absolute value to that of the harmonic mean. However, MTB=1 is worse than all comparable methods in terms of mean, standard deviation, and RMSE for log errors (Panel A). A very similar picture obtains for absolute percentage errors.

[Insert Tables 7 and 8 about here]

Tables 7 and 8 display the results for, respectively, the value-to-sales ratio and the price-earnings ratio. Instead of MTB=1 we use, respectively, Value=Sales and P/E=10 as *ad hoc* valuation methods. We do not continue to use MTB=1, because setting the market value of the firm equal to the book value would not mean using *less* information than the averaging methods in the table but using *different* information. Tables 7 and 8 show that all results for the market-to-book ratio continue to hold for the other two ratios. Note that the statistics in Tables 6 to 8 are not comparable across tables, because they refer to slightly different sets of firms. For instance, firms with negative earnings are included in the samples analyzed in Tables 6 and 7 while they have been excluded for the calculations shown in Table 8. Overall, the empirical analysis clearly corroborates our theoretical analysis and the simulations based on the lognormal distribution. This supports our strategy of using the lognormal model instead of more complicated distributional models that might fit the data better than the lognormal distribution.

5 The diversification discount: an application

In this section we apply our insights about alternative averaging procedures for multiples and the implications of choosing different error measures for the analysis of the diversification discount. The diversification discount is defined as the difference between the value of a diversified (multi-segment) firm and an imputed break-up value estimated as the sum of the values of each of its segments, where imputed segment values are calculated using a comparables approach, typically based on the value-to-sales ratio, the value-to-assets ratio, or Tobin's Q. We add this discussion because the large literature testifies to the continued interest in the correct measurement of the diversification discount. It is arguably the most prominent academic application of multiples, and it allows us to show that the methodological points raised in this paper are relevant for practitioners and researchers alike. Note that our concern here is only the correct measurement of the diversification discount, and we do not delve into what causes it or how it varies across countries or industries.[18]

The literature on the conglomerate discount estimates the imputed firm value $\widehat{MV}$ either with the arithmetic mean or with the median of comparable firms' financial ratios.[19] The harmonic mean or the geometric mean are never used. The literature has also not converged on the question how the conglomerate discount should be reported. Lang and Stulz (1994) and Servaes (1996) report the difference between the imputed ratio $\widehat{MV}/B$ and the conglomerate's ratio MV/B, where MV is the observed enterprise value, B is the multiple base (i.e. assets or sales), and $\widehat{MV}$ is the estimated break-up value obtained by adding the comparables-based valuations of all segments of the firm. This measure of the conglomerate discount corresponds to our percentage valuation errors above when the difference $\widehat{MV} - MV$ is scaled by B instead of MV. By contrast, Berger and Ofek (1995), Lins and Servaes (1999), and Schoar (2002) measure the excess value by $\ln(\widehat{MV}/MV)$, which corresponds to our definition of log valuation errors above. Villalonga (2004) uses both measures. Note that

[18] The survey by Martin and Sayrak (2003) provides a comprehensive review of the literature on diversification as well as a discussion of what causes the diversification discount.

[19] Lang and Stulz (1994) and Villalonga (2004) use the arithmetic mean. Berger and Ofek (1995), and Lins and Servaes (1999) use the median, and Servaes (1996) reports results for both methods.

with a procedure that overvalues the break-up value $\widehat{MV}$ more, we obtain a larger value for both measures and for the conglomerate discount. Conversely, a valuation procedure that undervalues $\widehat{MV}$ sufficiently strongly would lead us to report a diversification premium.

For our analysis of the conglomerate discount, we use Worldscope data, which provide us with a longer time-period for our analysis compared to our version of Compustat.[20] We proceed in the standard way described in the literature by identifying all multiple-segment firms on Worldscope domiciled in the United States.[21] We only consider companies with complete data for the first segment (SIC-code, sales, total assets), with company-wide sales and total assets both larger than \$1 million, with fiscal-year end in December, and with non-missing enterprise value, where enterprise value is defined as the market value of equity plus the book value of debt.[22] We exclude companies with a financial segment (SIC codes 6000 to 6999) and companies whose company-wide sales are not within 1% of the sum of the sales for all segments combined.[23] We obtain a sample with 4,756 firm-years for the period 1991-2003 and report results for the combined sample as well as for 2003 separately (438 firms). We match each segment to the members of its narrowest SIC-industry where we can identify at least 5 comparable single-segment companies. Our sample contains 16,397 single-segment firms (1,702 in 2003).

We report results for the asset multiplier and the sales multiplier in Table 9. Given the close relationship between asset multipliers and Tobin's Q we do not report results for Tobin's Q, which are very similar. The statistics in Table 9 are calculated in the same way as those in Tables 3 to 8. Following Berger and Ofek (1995), we delete all observations where $\ln(\widehat{MV}/MV)$ is larger than 1.386 or smaller than -1.386 before we calculate the statistics

[20] We have access to segment data from Compustat only for 1998-2004. For these years we obtain very similar results to those obtained with Worldscope.

[21] A firm is identified as a multi-segment firm if it has segments from at least two different 2-digit SIC groups (where headquarters with SIC code '9999' are disregarded).

[22] We follow the convention in this literature and use the market value of equity at the end of December. Note that this is different from our procedure in the last section, where we used the market value four months after fiscal-year end.

[23] When we check whether sales add up to total sales, we include headquarters segments with SIC code '9999.' If segment reporting is complete in this sense, we delete '9999' segments and rescale segment sales and segment assets such that they add up to company-wide sales and total assets, respectively. We delete all firms with negative sales or with negative assets in any of the remaining segments.

shown in Table 9. This deletes all firms where the estimated value $\widehat{MV}$ is more than 4 times (or less than 25% of) the market value MV.

In the previous tables, we interpreted a positive mean as a systematic overvaluation of target firms by the respective valuation method. Now, a positive mean is interpreted as a conglomerate discount: The sum of the individual segment values (calculated with a particular valuation method) is systematically higher than the market value of the conglomerate.

[Insert Table 9 about here]

Table 9 displays similar numbers to those reported in the literature: If we value segments with the arithmetic mean of the comparable firms' asset multipliers, we obtain an average conglomerate discount of 38.7% for the period from 1991 to 2003 (see Panel B). Villalonga (2004) reports a discount of 47% for this period. If we average comparable information with the median, we obtain a conglomerate discount of 14.5% with the asset multiplier, where Berger and Ofek (1995) obtain a discount of 12.2% for a sample that spans from 1986 to 1991. With the sales multiplier, our estimate (21.9%, see Panel A) is somewhat higher than that of Berger and Ofek (1995), who obtain only 9.7%.

Note that the imputed values based on medians or the geometric mean are similar to each other and much lower than estimates based on the arithmetic mean. This is consistent with our previous discussion: the median and the geometric mean always generate a lower estimate $\widehat{MV}$ of the imputed value of the firm and therefore also a lower estimate of the diversification discount (see Proposition 1). According to our results in Sections 2 and 3, the arithmetic mean is likely to overestimate the break-up value and therefore the conglomerate discount, whereas the geometric mean and the median will provide correct estimates. By comparison, the harmonic mean reports no diversification discount for the asset multiplier and a diversification *premium* of 3% for the sales multiplier.[24] The reason is that the harmonic mean undervalues $\widehat{MV}$ and compares the actual value of a firm to a benchmark that

[24] All estimates of the conglomerate discount in Table 9 increase slightly when conglomerates with sales less than $20m are excluded from the sample (not shown in the tables). The conglomerate premium found with the sales multiplier and the harmonic mean is then only 2%.

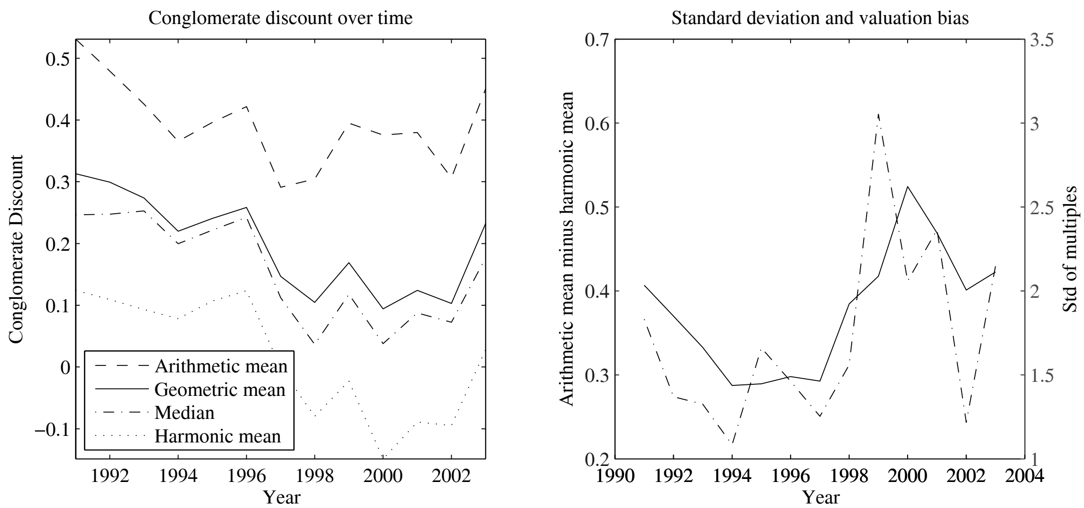

Figure 4: The left panel of this figure shows the conglomerate discount over time from 1991 to 2003. For each line in the figure the imputed (break-up) value of the firm is computed using a different multiple valuation method. The discount is calculated with the log error. Higher numbers are higher discounts, negative numbers correspond to a conglomerate premium. The right panel shows the difference between the conglomerate discount calculated with the arithmetic mean and the discount calculated with the harmonic mean (solid line, left scale) together with the average standard deviation of the comparable multiples (broken line, right scale). The correlation between the two series is 0.63.

is too low.[25]

It is instructive to see how the estimates of the diversification discount have changed over time depending on the averaging procedure chosen. The left panel of Figure 4 plots the results for our sample over time for the asset multiplier. The median and the geometric mean show that the conglomerate discount decreased from 25-30% in 1991 to about 10% in 1998 and the following years. The solid line in the right panel of Figure 4 shows the spread between arithmetic mean and harmonic mean, which increased from 30% in the mid-nineties to more than 50% in 2000. In that year, the conglomerate discount based on the arithmetic mean increases to almost 38% in 2000, while the conglomerate premium based on the harmonic mean rises to 15%. In fact, this illustrates the impact of rising standard deviations on the bias. In order to see this, the right panel of Figure 4 also displays the

[25] Observe that this reason for a conglomerate premium is completely unrelated to the conglomerate premium found in Villalonga (2004). Using a Compustat dataset similar to ours and a similar methodology she also finds a significant conglomerate discount.

average standard deviation of the comparable multiples used in the valuations[26] (broken line). The correlation between spread and volatility is 0.63 and it is significant at the 5% level. Due to the high dispersion of financial ratios during the "internet bubble" period, the arithmetic mean gives the wrong impression of an increased conglomerate discount in 1999-2001 that is not recorded by the median or the geometric mean. Conversely, the harmonic mean gives rise to a spurious conglomerate premium.

Our results suggest that those papers that reported a diversification discount on the basis of imputed values estimated with arithmetic means have greatly overstated the diversification discount. Those papers that use medians come up with more accurate (and much less spectacular) estimates for the discount. Moving to estimates based on the harmonic mean seems unwarranted.

[Insert Table 10 about here]

In Table 10, we report the conglomerate discount calculated with *percentage* valuation differences ($\widehat{MV}/MV-1$) rather than log differences ($\ln(\widehat{MV}/MV)$). A closely related (but not identical) method is used by Lang and Stulz (1994) and Servaes (1996), as argued at the beginning of this subsection. In line with our results from previous sections, conglomerate discount estimates from percentage differences are much larger than estimates from log differences. With the arithmetic mean one would obtain a discount of up to nearly 100%, i.e. the imputed value is nearly twice as large as the actual market value. This confirms our strong reservations about percentage differences. We do not believe that these numbers are good estimates of the conglomerate discount and report them for completeness only.

6 Discussion and conclusion

In this paper we investigate the bias and dispersion of different averaging procedures for multiples. We point out that in *any* sample, the arithmetic mean always results in higher

[26] For this, we first calculate the standard deviation of the comparable multiples for each segment that is valued. Then we compute the weighted average of these standard deviations across the segments of each firm in order to arrive at one volatility estimate for each firm. Finally, we average these volatility estimates across all firms in a given year.

valuations than the geometric mean, which in turn always predicts higher values than the harmonic mean. The differences between the valuations increase with the dispersion of values in the sample of comparable companies. Investment bankers and practitioners who use multiples to value acquisition targets or IPOs seem to have an intuitive grasp of the necessity to reduce the variation among comparables. They typically inspect the distribution and eliminate what appear to be outliers that are not representative of the industry and can be attributed to circumstances inapplicable to the target firm. Clearly, this ex post pruning of the sample is entirely *ad hoc*, but it reduces the variation in the set of comparable companies and makes the prediction less sensitive to potential biases of the averaging method. We take a different route here and analyze the bias and accuracy of valuation methods.

Our analysis shows that the answer to the question which averaging method leads to the best predictions depends crucially on the error measure chosen. We compare percentage errors with logarithmic errors. When percentage errors are used, the answer also depends on the dispersion within the set of comparable firms. In contrast, rankings of valuation methods obtained with log errors are robust to changes in dispersion. We reproduce the well-known fact that the arithmetic mean is positively biased and the harmonic mean is the least biased method with percentage errors. However, with log errors the harmonic mean is biased downward just as much as the arithmetic mean is biased upward, whereas the geometric mean is optimal in terms of bias and the bias of the median is negligible, though statistically significant. It is therefore critical to analyze the two error measures further.

Error measures are eventually based on loss functions which are inherently subjective and therefore beyond the scrutiny of normative analysis. So if the researcher or analyst who applies valuation procedures believes that the percentage error is the relevant error measure, then she will prefer the harmonic mean over all other averaging methods. However, this error measure implies a number of counterintuitive consequences:

- Ignoring all comparable information and setting the target firm value equal to its book value results in more precise and less biased forecasts than using the arithmetic mean or the median and becomes optimal for sufficiently dispersed samples.

- Ignoring all information altogether and setting the target firm value equal to \$1 turns out to be a reasonable valuation method when judged by percentage errors and becomes optimal if the dispersion of the sample is sufficiently large.

- As the variation among comparable firms increases, the bias of *all* averaging methods increases. Eventually, even the harmonic mean features a large positive bias.

The intuition for these results is that undervaluations measured with percentage errors are capped by -100% from below by virtue of limited liability, whereas overvaluations can and do become very large. Therefore, *ad hoc* methods that avoid extreme overvaluations and methods that are biased downward (like the harmonic mean) appear favorable with percentage errors. None of these counterintuitive results can be found for log errors, which penalize relative over- and undervaluations symmetrically.

On the basis of our theoretical and empirical results, we argue strongly in favor of using log errors as the appropriate error metric when comparing different valuation methods. As a consequence, we also argue in favor of the median and the geometric mean, and against the arithmetic as well as the harmonic mean. Given our results we find it therefore quite surprising that the geometric mean is the only standard averaging method that has not been widely adopted in the valuation literature. All distributions of multiples we investigate are skewed. The geometric mean effectively applies the arithmetic mean to the logarithms of multiples, thereby creating a symmetric distribution before averaging.

We apply our insights to the computation of the diversification discount as a prominent example for the use of comparables in financial research. Consistent with our previous argument we find that we overestimate the diversification discount if we average multiples with the arithmetic mean. Conversely, adopting the harmonic mean would – in our view erroneously – lead to estimates of the conglomerate discount around zero or even to a small conglomerate premium.

While our analysis has been conducted entirely on a small selection of widely used multiples, we argue that our results apply to other valuation methods as well, including those based on present value approaches. The asymmetric nature of percentage errors is independent of the valuation methodology, and therefore the implications of our analysis carry over to any valuation method. Many researchers (see the table in Appendix A) have staged "horse races" between multiples and present value-approaches, or between DCF and residual income. We conclude that to the extent that they are based on percentage valuation errors, these papers predispose their analysis to those valuation procedures that tend to undervalue companies while unduly criticizing approaches that tend to overvalue companies. Redoing these comparisons with logarithmic errors that properly take into account the asymmetric distribution of valuation errors is clearly on the agenda for future research on company valuation.

Ultimately, the error measure chosen must depend on the application in question. The objective function for a bidder in an auction for a company may be different from that of a security analyst who values a market traded company for investment purposes. Depending on risk aversion, degree of diversification, asymmetric information, and other considerations, practitioners will give different weights to small valuation errors versus large errors. Also, they wish to equate either equal dollar mispricings (the case for percentage errors) or equal relative mispricings (the case for log errors), or treat undervaluations and overvaluations differently altogether. Rigorous answers to these questions can only be obtained based on an axiomatic approach that relates decision rules to preferences and to the salient features of the application. We are not aware that such an approach has ever been pursued and believe that this will be a fertile area for future research.

Appendix

A Empirical Comparisons of Valuation Methods in the Literature

The following table summarizes 14 papers which empirically compare the accuracy of different valuation methods. The table shows the country and the period for which the study has been done, the error measure used in the comparison, and the multiple methods that are considered. If the study also considers present value methods, then we list them in the last column. Here DCF stands for “discounted cash-flows,” DDM for “dividend discount model,” and RIM for “residual income model.”

Paper	Country	Period(s)	Error measure(s)	Multiple methods	Discounting methods
Alford (1992)	USA	1978, 1982, 1986	Median of absolute percentage errors	Median	none
Baker & Ruback (1999)	USA	1995	Percentage errors	Harmonic mean	none
Beatty, Riffe & Thompson (1999)	USA	1980-1992	Percentage errors	Median; arithmetic, harmonic, and industry mean	none
Bhojraj & Lee (2002)	USA	1982-1998	Percentage errors, regression R^2	Harmonic mean	none
Cheng & McNamara (2000)	USA	1973-1992	Percentage errors	Median	none
Francis, Olsson & Oswald (2000)	USA	1989-1993	Percentage errors	none	DCF, DDM, RIM
Gilson, Hotchkiss & Ruback (2000)	USA	1984-1993	Log errors	Median	DCF
Herrmann & Richter (2003)	USA and Europe	1997-1999	Log errors	Median; arithmetic, harmonic, and geometric mean	none
Kaplan & Ruback (1995)	USA	1983-1989	Log errors	Median	DCF
Kim & Ritter (1999)	USA	1992-1993	Log errors	Median	none
Lie & Lie (2002)	USA	1998	Log errors	Median	none
Liu, Nissim & Thomas (2002a)	USA	1982-1999	Percentage errors	Harmonic mean, regression measures	none
Liu, Nissim & Thomas (2002b)	10 countries	1976-2001	Percentage errors	Harmonic mean	none
Penman & Sougiannis (1998)	USA	1973-1992	Percentage errors	none	DCF, DDM, RIM

B Proof of Propositions

B.1 Proof of Proposition 1

(i) From Jensen's inequality we have:

$$\ln\left(\bar{x}_j^A\right) = \ln\left(\frac{1}{n}\sum_{i=1}^{n} x_i\right) > \frac{1}{n}\sum_{i=1}^{n} \ln\left(x_i\right) = \ln\left(\bar{x}_j^G\right). \tag{8}$$

Similarly, we can apply Jensen's inequality to the rewritten harmonic mean:

$$\ln\left(\bar{x}_j^H\right) = -\ln\left(\frac{1}{n}\sum_{i=1}^{n} \frac{1}{x_i}\right) < -\frac{1}{n}\sum_{i=1}^{n} \ln\left(1/x_i\right) = \frac{1}{n}\sum_{i=1}^{n} \ln\left(x_i\right) = \ln\left(\bar{x}_j^G\right) \ .$$

As the logarithm is a monotonic transformation, we obtain $\bar{x}_j^H < \bar{x}_j^G < \bar{x}_j^A$.

(ii) Let M denote the median and μ the mean of the distribution of the ratios x_i. As the distribution is symmetric, we have $M = \mu$. The law of large numbers implies that sample moments converge to population moments, so $\bar{x}_j^M - \bar{x}_j^A \rightarrow M - \mu = 0$ as $n \rightarrow \infty$.

(iii) Denote the parameters of the lognormal distribution from which the x_i have been drawn by μ and σ^2. Then, the median of the distribution is $M = \exp\{\mu\}$ and $E\left(\ln\left(x_i\right)\right) = \mu$, so that $\bar{x}_j^M - \bar{x}_j^G = \bar{x}_j^M - \exp\left\{\frac{1}{n}\sum_{i=1}^{n} \ln\left(x_i\right)\right\} \rightarrow M - \exp(\mu) = 0$ as $n \rightarrow \infty$ by the same argument as in (ii).

B.2 Proof of Proposition 2

We only need to show that $E\left(e_{\log}^G\right) = 0$, $E\left(e_{pct}^G\right) > 0$, and $E\left(e_{\log}^H\right) = -E\left(e_{\log}^A\right)$. The remaining statements of the proposition then follow immediately from Proposition 1.

$>$From (4) we have:

$$\ln\left(\ \bar{x}^G\right) = \frac{1}{n}\sum_{i=1}^{n} \ln\left(x_i\right) \ ,$$

so $\ln\left(\bar{x}^G\right)$ is distributed normal with mean μ and variance σ^2/n. Hence, $e^G_{\log} = \ln\left(\bar{x}^G\right) - \ln(x_j)$ is distributed normal with mean zero and variance $\frac{1+n}{n}\sigma^2$, so the geometric mean leads to unbiased estimates in terms of logarithmic errors. As a consequence, $1 + e^G_{pct} = \exp\left(e^G_{\log}\right)$ is distributed lognormal with parameters 0 and $\frac{1+n}{n}\sigma^2$, so we obtain:

$$E\left(e^G_{pct}\right) = \exp\left(\frac{1+n}{2n}\sigma^2\right) - 1 > 0 \ ,$$

as long as $\sigma^2 > 0$. This shows that the geometric mean leads to biased estimates in terms of percentage errors.

Showing that $E\left(e^H_{\log}\right) = -E\left(e^A_{\log}\right)$ requires a little more work: First note that

$$E\left(\ln\left(\bar{x}^H\right)\right) = -E\left(\ln\left(\frac{1}{n}\sum_{i=1}^{n}\exp\left(-\ln x_i\right)\right)\right) = -E\left(\ln\left(\frac{1}{n}\sum_{i=1}^{n}\exp\left(-u_i\right)\right)\right), \qquad (9)$$

where $u_i = \ln x_i$, which is distributed normal with expectation μ and variance σ^2. We expand this expression and perform the substitution $v_i = -u_i + 2\mu$ for all $i = 1...n$.

$$\begin{aligned}
&E\left(\ln\left(\bar{x}^H\right)\right) = \\
&= -\left[\sigma\sqrt{2\pi}\right]^{-n}\int_{-\infty}^{\infty}...\int_{-\infty}^{\infty}\ln\left(\frac{1}{n}\sum_{i=1}^{n}\exp\left(-u_i\right)\right)\prod_{i=1}^{n}\exp\left\{-\frac{1}{2}\left(\frac{u_i-\mu}{\sigma}\right)^2\right\}du_1...du_n \\
&= -\left[\sigma\sqrt{2\pi}\right]^{-n}\int_{-\infty}^{\infty}...\int_{-\infty}^{\infty}\ln\left(\frac{1}{n}\sum_{i=1}^{n}\exp\left(v_i-2\mu\right)\right)\prod_{i=1}^{n}\exp\left\{-\frac{1}{2}\left(\frac{-v_i+\mu}{\sigma}\right)^2\right\}dv_1...dv_n \\
&= 2\mu-\left[\sigma\sqrt{2\pi}\right]^{-n}\int_{-\infty}^{\infty}...\int_{-\infty}^{\infty}\ln\left(\frac{1}{n}\sum_{i=1}^{n}\exp\left(v_i\right)\right)\prod_{i=1}^{n}\exp\left\{-\frac{1}{2}\left(\frac{v_i-\mu}{\sigma}\right)^2\right\}dv_1...dv_n \\
&= 2\mu - E\left(\ln\left(\frac{1}{n}\sum\nolimits_{i=1}^{n}\exp u_i\right)\right) = 2\mu - E\left(\ln\left(\bar{x}^A\right)\right)
\end{aligned}$$

The second line rewrites (9) more explicitly and the third line applies the transformation $u_i = -v_i + 2\mu$. Here, we use the fact that $\frac{dv_i}{du_i} = -1$, which cancels with the factor -1 caused by the necessary transformation of the integration limits. The fourth line follows upon rearranging, and the fifth line rewrites the same expression using the expectations operator. Observe that v_i is also distributed normal with mean μ and variance σ^2. Hence, $E\left(\ln\left(\bar{x}^H\right)\right) - \mu = -\left[E\left(\ln\left(\bar{x}^A\right)\right) - \mu\right]$. As $E\left(\ln(x_i)\right) = \mu$, this implies $E\left(e^H_{\text{log}}\right) = -E\left(e^A_{\text{log}}\right)$ from the definition of log errors (7).

References

[1] Alford, Andrew W., 1992, *The Effect of the Set of Comparable Firms on the Accuracy of the Price-Earnings Valuation Method,* Journal of Accounting Research 30, no. 1 , pp. 94-109

[2] Baker, Malcolm, and Richard S. Ruback, 1999, *Estimating Industry Multiples,* Mimeo, Harvard University, (June)

[3] Beatty, Randolph P.; Susan M. Riffe, and Rex Thompson, 1999, *The Method of Comparables and Tax Court Valuations of Private Firms: An Empirical Investigation,* Accounting Horizons 13, no. 3 , pp. 177-199

[4] Berger, Philip G., and Eli Ofek, 1995, *Diversification's Effect on Firm Value,* Journal of Financial Economics 37, pp. 39-65

[5] Bhojraj, Sanjeev, and Charles M. C. Lee, 2001, *Who Is My Peer? A Valuation-Based Approach to the Selection of Comparable Firms,* Journal of Accounting Research 40, no. 2 (September), pp. 407-439

[6] Cheng, C. S. Agnes, and Ray McNamara, 2000, *The Valuation Accuracy of the Price-Earnings and Price-Book Benchmark Valuation Methods,* Review of Quantitative Finance and Accounting 15, pp. 349-370

[7] D'Agostino, Ralph B.; Michael A. Stephens, 1986, *Goodness-of-fit techniques*, New York (Marcel Dekker)

[8] Francis, Jennifer; Per Olsson, and Dennis R. Oswald, 2000, *Comparing the Accuracy and Explainability of Dividend, Free Cash Flow and Abnormal Earnings Equity Value Estimates,* Journal of Accounting Research 38, pp. 45-70

[9] Gilson, Stuart C.; Edith S. Hotchkiss, and Richard S. Ruback, 2000, *Valuation of Bankrupt Firms,* Review of Financial Studies 13, no. 1 , pp. 43-74

[10] Herrmann, Volker, and Frank Richter, 2003, *Pricing With Performance-Controlled Multiples,* Schmalenbach Business Review 55, (July), pp. 194-219

[11] Kaplan, Steven N., and Richard S. Ruback, 1995, *The Valuation of Cash Flow Forecasts: An Empirical Analysis,* Journal of Finance 50, no. 4 (September), pp. 1059-1093

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

[13] Lang, Larry, and Rene Stulz, 1994, *Tobin's Q, Corporate Diversification and Firm Performance,* Journal of Political Economy 102, pp. 1248-1280

[14] Lie, Erik, and Heidi J. Lie, 2002, *Multiples Used to Estimate Corporate Value,* Financial Analysts Journal, pp. 44-54

[15] Lins, Karl, and Henri Servaes, 1999, *International Evidence on the Value of Corporate Diversification,* Journal of Finance 54, no. 6 (December), pp. 2215-2239

[16] Liu, Jing; Doron Nissim, and Jacob Thomas, 2002a, *Equity Valuation Using Multiples,* Journal of Accounting Research 40, no. 1 (March), pp. 135-172

[17] Liu, Jing; Doron Nissim, and Jacob Thomas, 2002b, *International Equity Valuation Using Multiples,* Mimeo, University of California at Los Angeles

[18] Martin, John D., and Akin Sayrak, 2003, *Corporate Diversification and Shareholder Value,* Journal of Corporate Finance 9, pp. 37-57

[19] Penman, Stephen H., and Theodore Sougiannis, 1998, *A Comparison of Dividend, Cash Flow, and Earnings Approaches to Equity Valuation,* Contemporary Accounting Research 15, no. 3 (Fall), pp. 343-383

[20] Schoar, Antoinette S., 2002, *Effects of Corporate Diversification on Productivity,* Journal of Finance 57, no. 6 (December), pp. 2379-2403

[21] Servaes, Henri, 1996, *The Value of Diversification During the Conglomerate Merger Wave,* Journal of Finance 51, no. 4 (September), pp. 1201-1225

[22] Villalonga, Belen, 2004, *Diversification Discount or Premium? New Evidence From BITS Establishment-Level Data,* Journal of Finance 59, no. 2 (April), pp. 479-506

Table 1: Descriptive Statistics for Multiples

This table displays descriptive statistics (number of observations, mean, median, standard deviation, skewness, excess kurtosis, and the 10% and 90% quantile) for the distributions of the market-to-book ratio, the value-to-sales ratio, and the price-earnings ratio for the pooled sample from 1994 to 2003. The lower part of the table shows the statistics for the natural logarithms of these ratios. For each multiple and each year, the multiples have been winzorized at the 2.5% and 97.5% quantile.

Multiple	# obs.	Mean	Std. dev.	Skewness	Kurtosis	P10	Median	P90
market-to-book	47,614	3.031	3.513	3.295	14.717	0.682	1.867	6.601
value-to-sales	51,899	3.583	6.244	5.959	52.318	0.392	1.634	8.027
price-earnings	33,753	29.593	40.133	3.692	15.365	7.520	17.073	57.317
log(market-to-book)	47,614	0.691	0.882	0.280	0.097	-0.382	0.624	1.887
log(value-to-sales)	51,899	0.549	1.170	0.260	-0.291	-0.937	0.491	2.083
log(price-earnings)	33,753	2.945	0.852	0.653	0.980	2.018	2.837	4.049

Table 2: Tests for Normality

This tables shows the test statistics of three tests for normality (Kolmogorov-Smirnov, Cramér-von Mises, and Anderson-Darling) applied to the pooled sample (1994-2003) of three different financial ratios and their logarithmic transformations. All observations have been standardized by deducting the industry-year mean and then dividing by the industry-year standard deviation, where the industry is given by the two-digit SIC code. For the Kolmogorov-Smirnov test, the table displays the usual test statistic multiplied by the square root of the number of observations in order to make comparisons across samples meaningful. For the thus transformed Kolmogorov-Smirnov test statistic, the 1% critical value is 1.035. For the Cramér-von Mises and the Anderson-Darling test the 1% critical value are 0.179 and 1.035, respectively. Critial values have been obtained from D'Agnostino and Stephens (1986), p.123.

Multiple	Test statistics		
	Kolmogorov-Smirnov	Cramer-von Mises	Anderson-Darling
market-to-book	36.60	537.85	3,006.19
value-to-sales	43.53	723.11	3,871.00
price-earnings	43.70	708.97	3,720.94
log(market-to-book)	3.76	4.84	32.52
log(value-to-sales)	3.88	5.36	36.30
log(price-earnings)	12.66	51.50	287.46

Table 3: Simulated error distributions for industries with 40 companies

This table displays descriptive statistics of the simulated valuation errors from six valuation methods based on the market-to-book ratio. It shows the mean, median, root mean squared error, the standard deviation, skewness, and the proportion of observations larger than or equal to +100%. In each of the 100,000 runs, we simulate an industry with 40 companies and value each of these companies using comparable information from the remaining 39 firms. For the methods 'arithmetic mean,' 'median,' 'geometric mean,' and 'harmonic mean,' the 39 market-to-book ratios are averaged with the respective method and the result is multiplied by the target firm's book value to arrive at a forecast of the target firm's market value. The method 'MTB=1' sets the target firm's market value equal to its book value, and the method 'Value=$1' sets the target firm's market value equal to $1. For the simulation we assume that market-to-book ratios are lognormally distributed with mean 0.597 and standard deviation 0.801, and that the book value is lognormally distributed with mean 18.312 and standard deviation 1.828. Panel A shows the results for log errors, Panel B for percentage errors, and panel C for absolute percentage errors.

Panel A: Log errors

Method	Mean	Median	RMSE	Std. dev.	Skewness	Prop>=100%
Arithmetic mean	0.308	0.309	0.872	0.815	-0.001	0.198
Median	0.000	0.000	0.817	0.817	-0.001	0.110
Geometric mean	0.000	0.000	0.811	0.811	-0.002	0.109
Harmonic mean	-0.308	-0.309	0.872	0.815	-0.003	0.054
MTB=1	-0.597	-0.596	0.999	0.801	-0.002	0.023
Value=$1	-18.909	-18.910	19.014	1.995	0.000	0.000

Panel B: Percentage errors

Method	Mean	Median	RMSE	Std. dev.	Skewness	Prop>=100%
Arithmetic mean	0.899	0.363	2.048	1.840	3.728	0.319
Median	0.395	0.000	1.412	1.355	3.741	0.198
Geometric mean	0.389	0.000	1.392	1.337	3.691	0.196
Harmonic mean	0.023	-0.266	0.989	0.989	3.715	0.109
MTB=1	-0.241	-0.449	0.757	0.718	3.607	0.054
Value=$1	-1.000	-1.000	1.000	0.000	87.845	0.000

Panel C: Absolute percentage errors

Method	Mean	Median	RMSE	Std. dev.	Skewness	Prop>=100%
Arithmetic mean	1.162	0.584	2.048	1.686	4.522	0.319
Median	0.818	0.508	1.412	1.151	5.181	0.198
Geometric mean	0.809	0.505	1.392	1.133	5.132	0.196
Harmonic mean	0.640	0.508	0.989	0.755	5.925	0.109
MTB=1	0.577	0.549	0.757	0.490	5.959	0.054
Value=$1	1.000	1.000	1.000	0.000	-87.845	0.000

Table 4: Simulated error distributions for industries with 5 companies

This table displays descriptive statistics of the simulated valuation errors from six valuation methods based on the market-to-book ratio. It shows the mean, median, root mean squared error, the standard deviation, skewness, and the proportion of observations larger than or equal to +100%. In each of the 100,000 runs, we simulate an industry with 5 companies and value each of these companies using comparable information from the remaining 4 firms. For the methods 'arithmetic mean,' 'median,' 'geometric mean,' and 'harmonic mean,' the 4 market-to-book ratios are averaged with the respective method and the result is multiplied by the target firm's book value to arrive at a forecast of the target firm's market value. The method 'MTB=1' sets the target firm's market value equal to its book value, and the method 'Value=$1' sets the target firm's market value equal to $1. For the simulation we assume that market-to-book ratios are lognormally distributed with mean 0.597 and standard deviation 0.801, and that the book value is lognormally distributed with mean 18.312 and standard deviation 1.828. Panel A shows the results for log errors, and Panel B for percentage errors.

Panel A: Log errors

Method	Mean	Median	RMSE	Std. dev.	Skewness	Prop>=100%
Arithmetic mean	0.223	0.222	0.939	0.912	0.007	0.197
Median	0.046	0.046	0.916	0.915	-0.006	0.148
Geometric mean	0.000	-0.001	0.895	0.895	-0.003	0.132
Harmonic mean	-0.223	-0.223	0.939	0.912	-0.011	0.090
MTB=1	-0.595	-0.595	0.998	0.801	0.000	0.023
Value=$1	-18.903	-18.905	19.008	1.995	0.005	0.000

Panel B: Percentage errors

Method	Mean	Median	RMSE	Std. dev.	Skewness	Prop>=100%
Arithmetic mean	0.897	0.249	2.340	2.162	4.672	0.302
Median	0.589	0.048	1.902	1.809	4.655	0.239
Geometric mean	0.492	-0.001	1.714	1.642	4.450	0.219
Harmonic mean	0.211	-0.200	1.381	1.365	4.645	0.157
MTB=1	-0.240	-0.448	0.757	0.718	3.554	0.054
Value=$1	-1.000	-1.000	1.000	0.000	59.770	0.000

Table 5: The influence of an increase of dispersion on the simulated error distributions

This table displays descriptive statistics of the simulated valuation errors from six valuation methods based on the market-to-book ratio. It shows the mean, median, root mean squared error, the standard deviation, skewness, and the proportion of observations larger than or equal to +100%. In each of the 100,000 runs, we simulate an industry with 40 companies and value each of these companies using comparable information from the remaining 39 firms. For the methods 'arithmetic mean,' 'median,' 'geometric mean,' and 'harmonic mean,' the 39 market-to-book ratios are averaged with the respective method and the result is multiplied by the target firm's book value to arrive at a forecast of the target firm's market value. The method 'MTB=1' sets the target firm's market value equal to its book value, and the method 'Value=$1' sets the target firm's market value equal to $1. For the simulation we assume that market-to-book ratios are lognormally distributed with mean 0.597 and standard deviation 1.202, and that the book value is lognormally distributed with mean 18.312 and standard deviation 1.828. Panel A shows the results for log errors, and Panel B for percentage errors.

Panel A: Log errors

Method	Mean	Median	RMSE	Std. dev.	Skewness	Prop>=100%
Arithmetic mean	0.682	0.682	1.405	1.228	0.003	0.398
Median	0.000	0.000	1.223	1.223	0.000	0.207
Geometric mean	0.000	0.001	1.214	1.214	-0.001	0.205
Harmonic mean	-0.682	-0.682	1.405	1.228	-0.004	0.085
MTB=1	-0.609	-0.609	1.345	1.199	0.000	0.090
Value=$1	-18.766	-18.766	18.905	2.295	-0.001	0.000

Panel B: Percentage errors

Method	Mean	Median	RMSE	Std. dev.	Skewness	Prop>=100%
Arithmetic mean	3.209	0.978	8.529	7.903	11.483	0.496
Median	1.110	0.000	4.064	3.909	11.080	0.285
Geometric mean	1.089	0.001	3.973	3.821	10.863	0.284
Harmonic mean	0.073	-0.494	2.001	2.000	11.352	0.131
MTB=1	0.115	-0.456	1.999	1.995	10.720	0.139
Value=$1	-1.000	-1.000	1.000	0.000	466.995	0.000

Table 6: Empirical error distributions for valuations based on the market-to-book ratio

This table displays descriptive statistics of the valuation errors from six valuation methods based on the market-to-book ratio. It is calculated from 47,614 firm-year observations from 1994 to 2003, and shows the mean, median, root mean squared error, the standard deviation, skewness, the proportion of observations larger than or equal to +100%, and the t-statistic of the two sided t-test that the mean equals zero. For the methods 'arithmetic mean,' 'median,' 'geometric mean,' and 'harmonic mean,' the industry peer group market-to-book ratios are averaged with the respective method and the result is multiplied by the target firm's book value to arrive at a forecast of the target firm's market value. The method 'MTB=1' sets the target firm's market value of equity equal to its book value, and the method 'Value=$1' sets the target firm's market value of equity equal to $1. Panel A shows the results for log errors, Panel B for percentage errors, and Panel C for absolute percentage errors.

Panel A: Log errors

Method	Mean	Median	RMSE	Std. dev.	Skewness	Prop>=100%	T-test
Arithmetic mean	0.281	0.248	0.940	0.897	-0.044	0.188	68.39
Median	-0.019	0.001	0.892	0.891	-0.224	0.106	-4.64
Geometric mean	-0.004	0.016	0.881	0.881	-0.235	0.108	-1.05
Harmonic mean	-0.274	-0.210	0.933	0.892	-0.407	0.058	-67.12
MTB=1	-0.694	-0.624	1.183	0.958	-0.363	0.030	-158.05
Value=$1	-18.790	-18.679	18.921	2.224	-0.248	0.000	-1843.42

Panel B: Percentage errors

Method	Mean	Median	RMSE	Std. dev.	Skewness	Prop>=100%	T-test
Arithmetic mean	1.103	0.281	6.756	6.665	76.713	0.290	36.12
Median	0.529	0.001	5.334	5.308	94.706	0.180	21.74
Geometric mean	0.531	0.016	5.151	5.124	92.111	0.181	22.63
Harmonic mean	0.158	-0.189	4.052	4.049	100.889	0.109	8.54
MTB=1	-0.176	-0.464	3.867	3.863	107.490	0.055	-9.94
Value=$1	-1.000	-1.000	1.000	0.000	26.609	0.000	-9.7E+08

Panel C: Absolute percentage errors

Method	Mean	Median	RMSE	Std. dev.	Skewness	Prop>=100%	T-test
Arithmetic mean	1.371	0.536	6.756	6.615	78.343	0.290	45.22
Median	0.940	0.461	5.334	5.250	97.607	0.180	39.09
Geometric mean	0.933	0.462	5.151	5.066	95.059	0.181	40.18
Harmonic mean	0.721	0.457	4.052	3.988	105.207	0.109	39.43
MTB=1	0.692	0.555	3.867	3.805	111.825	0.055	39.67
Value=$1	1.000	1.000	1.000	0.000	-26.609	0.000	9.7E+08

Table 7: Empirical error distributions for valuations based on the value-to-sales ratio

This table displays descriptive statistics of the valuation errors from six valuation methods based on the value-to-sales ratio. It is calculated from 51,899 firm-year observations from 1994 to 2003, and shows the mean, median, root mean squared error, the standard deviation, skewness, the proportion of observations larger than or equal to +100%, and the t-statistic of the two sided t-test that the mean equals zero. For the methods 'arithmetic mean,' 'median,' 'geometric mean,' and 'harmonic mean,' the industry peer group market-to-book ratios are averaged with the respective method and the result is multiplied by the target firm's book value to arrive at a forecast of the target firm's market value. The method 'Value=Sales' sets the target firm's enterprise value equal to its sales, and the method 'Value=$1' sets the target firm's enterprise value equal to $1. Panel A shows the results for log errors, Panel B for percentage errors, and Panel C for absolute percentage errors.

Panel A: Log errors

Method	Mean	Median	RMSE	Std. dev.	Skewness	Prop>=100%	T-test
Arithmetic mean	0.385	0.327	1.104	1.034	0.000	0.243	84.88
Median	-0.024	0.004	0.996	0.996	-0.394	0.122	-5.56
Geometric mean	0.000	0.023	0.988	0.988	-0.396	0.129	0.03
Harmonic mean	-0.343	-0.248	1.068	1.011	-0.710	0.064	-77.22
Value=Sales	-0.554	-0.491	1.369	1.252	-0.380	0.089	-100.88
Value=$1	-19.009	-18.893	19.143	2.260	-0.268	0.000	-1916.42

Panel B: Percentage errors

Method	Mean	Median	RMSE	Std. dev.	Skewness	Prop>=100%	T-test
Arithmetic mean	1.711	0.386	7.073	6.863	26.575	0.345	56.79
Median	0.638	0.004	3.959	3.907	39.860	0.199	37.22
Geometric mean	0.659	0.023	3.828	3.771	36.241	0.205	39.80
Harmonic mean	0.148	-0.220	2.466	2.461	40.817	0.115	13.71
Value=Sales	0.211	-0.388	3.495	3.489	43.830	0.149	13.76
Value=$1	-1.000	-1.000	1.000	0.000	119.342	0.000	-7.4E+08

Panel C: Absolute percentage errors

Method	Mean	Median	RMSE	Std. dev.	Skewness	Prop>=100%	T-test
Arithmetic mean	1.974	0.626	7.073	6.792	27.310	0.345	66.22
Median	1.068	0.492	3.959	3.812	42.585	0.199	63.81
Geometric mean	1.076	0.497	3.828	3.674	38.864	0.205	66.70
Harmonic mean	0.755	0.485	2.466	2.348	46.280	0.115	73.22
Value=Sales	0.998	0.683	3.495	3.350	48.815	0.149	67.85
Value=$1	1.000	1.000	1.000	0.000	-119.342	0.000	7.4E+08

Table 8: Empirical error distributions for valuations based on the price-earnings ratio

This table displays descriptive statistics of the valuation errors from six valuation methods based on the price-earnings ratio. It is calculated from 33,753 firm-year observations from 1994 to 2003, and shows the mean, median, root mean squared error, the standard deviation, skewness, the proportion of observations larger than or equal to +100%, and the t-statistic of the two sided t-test that the mean equals zero. For the methods 'arithmetic mean,' 'median,' 'geometric mean,' and 'harmonic mean,' the industry peer group market-to-book ratios are averaged with the respective method and the result is multiplied by the target firm's book value to arrive at a forecast of the target firm's market value. The method 'P/E=10' sets the target firm's market value of equity equal to ten times its net income, and the method 'Value=$1' sets the target firm's market value of equity equal to $1. Panel A shows the results for log errors, Panel B for percentage errors, and Panel C for absolute percentage errors.

Panel A: Log errors

Method	Mean	Median	RMSE	Std. dev.	Skewness	Prop>=100%	T-test
Arithmetic mean	0.307	0.305	1.041	0.995	-0.297	0.179	56.73
Median	-0.052	0.005	0.969	0.967	-0.496	0.081	-9.91
Geometric mean	-0.002	0.056	0.965	0.965	-0.519	0.088	-0.29
Harmonic mean	-0.256	-0.147	1.008	0.975	-0.713	0.052	-48.30
P/E=10	-0.646	-0.535	1.197	1.008	-0.742	0.029	-117.74
Value=$1	-19.200	-19.150	19.326	2.205	-0.142	0.000	-1599.11

Panel B: Percentage errors

Method	Mean	Median	RMSE	Std. dev.	Skewness	Prop>=100%	T-test
Arithmetic mean	1.823	0.357	16.935	16.837	32.569	0.288	19.89
Median	0.883	0.005	11.627	11.594	38.973	0.137	13.99
Geometric mean	0.948	0.057	11.405	11.366	35.796	0.154	15.32
Harmonic mean	0.466	-0.137	8.429	8.416	38.928	0.089	10.17
P/E=10	0.039	-0.414	6.590	6.590	41.377	0.047	1.08
Value=$1	-1.000	-1.000	1.000	0.000	31.522	0.000	-6.8E+08

Panel C: Absolute percentage errors

Method	Mean	Median	RMSE	Std. dev.	Skewness	Prop>=100%	T-test
Arithmetic mean	2.061	0.557	16.935	16.810	32.687	0.288	22.52
Median	1.270	0.383	11.627	11.558	39.238	0.137	20.18
Geometric mean	1.308	0.403	11.405	11.330	36.041	0.154	21.20
Harmonic mean	0.979	0.385	8.429	8.372	39.365	0.089	21.48
P/E=10	0.857	0.488	6.590	6.534	42.071	0.047	24.09
Value=$1	1.000	1.000	1.000	0.000	-31.522	0.000	6.8E+08

Table 9:

Estimates of the conglomerate discount with log errors

This table displays estimates of the conglomerate discount in the USA obtained with log errors for two different multiple valuation methods. Higher numbers are higher discounts, negative numbers correspond to a conglomerate premium. Panel A show the estimates when segments are valued based on the sales multiple, while Panel B show estimates based on the asset multiple. The left part of the table shows the mean, median and the standard deviation of the discount estimates for 438 firms from 2003. The right part of the table shows these statistics for 4,756 firms from the pooled sample from 1991 to 2003.

Panel A: Conglomerate discount measured with sales multiples

Method	2003			1991-2003		
	Mean	Median	Std. dev.	Mean	Median	Std. dev.
Arithmetic mean	0.532	0.622	0.593	0.483	0.542	0.564
Median	0.254	0.236	0.565	0.219	0.240	0.583
Geometric mean	0.309	0.311	0.563	0.249	0.283	0.575
Harmonic mean	-0.002	-0.010	0.589	-0.030	-0.029	0.591

Panel B: Conglomerate discount measured with asset multiples

Method	2003			1991-2003		
	Mean	Median	Std. dev.	Mean	Median	Std. dev.
Arithmetic mean	0.451	0.489	0.517	0.387	0.407	0.505
Median	0.177	0.206	0.491	0.145	0.155	0.503
Geometric mean	0.232	0.266	0.499	0.186	0.199	0.499
Harmonic mean	0.028	0.055	0.506	0.002	0.020	0.501

Table 10:

Estimates of the conglomerate discount with percentage errors

This table displays estimates of the conglomerate discount in the USA obtained with percentage errors for two different multiples valuation methods. Higher numbers are higher discounts, negative numbers correspond to a conglomerate premium. Panel A show the estimates when segments are valued based on the sales multiple, while Panel B show estimates based on the asset multiple. The left part of the table shows the mean, median and the standard deviation of the discount estimates for 438 firms from 2003. The right part of the table shows these statistics for 4,756 firms from the pooled sample from 1991 to 2003.

Panel A: Conglomerate discount measured with sales multiples

Method	2003			1991-2003		
	Mean	Median	Std. dev.	Mean	Median	Std. dev.
Arithmetic mean	0.969	0.862	0.939	0.867	0.719	0.928
Median	0.503	0.266	0.840	0.462	0.272	0.826
Geometric mean	0.582	0.365	0.859	0.498	0.327	0.824
Harmonic mean	0.188	-0.010	0.739	0.155	-0.029	0.717

Panel B: Conglomerate discount measured with asset multiples

Method	2003			1991-2003		
	Mean	Median	Std. dev.	Mean	Median	Std. dev.
Arithmetic mean	0.769	0.630	0.826	0.659	0.503	0.790
Median	0.336	0.229	0.630	0.305	0.167	0.656
Geometric mean	0.414	0.304	0.663	0.357	0.220	0.667
Harmonic mean	0.160	0.057	0.576	0.132	0.020	0.578

Publications in the Report Series Research* in Management

ERIM Research Program: "Finance and Accounting"

2006

Valuation Biases, Error Measures, and the Conglomerate Discount
Ingolf Dittmann and Ernst Maug
ERS-2006-011-F&A

---

* A complete overview of the ERIM Report Series Research in Management:
https://ep.eur.nl/handle/1765/1

ERIM Research Programs:
LIS Business Processes, Logistics and Information Systems
ORG Organizing for Performance
MKT Marketing
F&A Finance and Accounting
STR Strategy and Entrepreneurship