

# Accepted Manuscript

## Measuring Firm Size in Empirical Corporate Finance

Chongyu Dang , Zhichuan (Frank) Li , Chen Yang

PII: S0378-4266(17)30220-0  
DOI: [10.1016/j.jbankfin.2017.09.006](https://doi.org/10.1016/j.jbankfin.2017.09.006)  
Reference: JBF 5210

To appear in: *Journal of Banking and Finance*

Received date: 21 August 2016  
Revised date: 4 September 2017  
Accepted date: 12 September 2017

Please cite this article as: Chongyu Dang , Zhichuan (Frank) Li , Chen Yang , Measuring Firm Size in Empirical Corporate Finance, *Journal of Banking and Finance* (2017), doi: [10.1016/j.jbankfin.2017.09.006](https://doi.org/10.1016/j.jbankfin.2017.09.006)

This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.



# Measuring Firm Size in Empirical Corporate Finance

Chongyu Dang, Zhichuan (Frank) Li, Chen Yang\*

This Version: September 12, 2017

## Abstract

In empirical corporate finance, firm size is commonly used as an important, fundamental firm characteristic. However, no research comprehensively assesses the sensitivity of empirical results in corporate finance to different measures of firm size. This paper fills this hole by providing empirical evidence for a  $\delta$   $\circ$   $g$   $c$   $u$   $w$   $t$   $g$   $\circ$   $g$   $p$   $v$   $"$   $g$   $h$   $g$   $e$   $v$   $\ddot{o}$   $"$   $k$   $p$   $"$   $t$   $h$   $e$   $\delta$   $u$   $k$   $|$   $g$   $"$   $g$   $h$   $g$   $e$   $v$   $\ddot{o}$   $\emptyset$  In particular, we examine the influences of employing different proxies (total assets, total sales, and market capitalization) of firm size in 20 prominent areas in empirical corporate finance research. We highlight several empirical implications. First, in most areas of corporate finance the coefficients of firm size measures are robust in sign and statistical significance. Second, the coefficients on regressors other than firm size often change sign and significance when different size measures are used. Unfortunately, this suggests that some previous studies are not robust to different firm size proxies. Third, the goodness of fit measured by R-squared also varies with different size measures, suggesting that some measures are more relevant than others in different situations. Fourth, different proxies  $e$   $c$   $r$   $v$   $w$   $t$   $g$   $"$   $f$   $k$   $h$   $g$   $t$   $g$   $p$   $v$   $"$   $c$   $u$   $r$   $g$   $e$   $v$   $"$   $q$   $h$   $"$   $\delta$   $h$   $k$   $t$   $\circ$   $"$   $u$   $k$   $|$   $g$   $\ddot{o}$ , and thus have different implications. Therefore, the choice of size measures needs both theoretical and empirical justification. Finally, our empirical assessment provides guidance to empirical corporate finance researchers who must use firm size measures in their work.

*JEL Classifications:* G3, G30, G31, G32, G34, G35, C23, C58, J31, J33.

*Key Words:* Firm size measures; Total assets; Total sales; Market capitalization; Empirical corporate finance.

---

\* We thank Geert Bekaert (the editor) and two anonymous reviewers for their valuable comments. For helpful insights, we are also grateful to Steve Foerster, Craig Dunbar, Stephannie Larocque, Stephen Sapp, Shahbaz Sheikh, Jeff Coles, Michael Schill, Susan Christoffersen, Francesca Cornelli, Pedro Matos, Heitor Almeida, Simi Kedia, Patrick Akey, Philip Strahan, Michael King, Walid Busaba, George Athanassakos, and seminar participants at Ivey Business School, the annual meeting of International Finance and Banking Society (IFABS) 2015, Hangzhou, China, and the annual meeting of Canadian Law and Economics Association (CLEA) 2015, University of Toronto, Canada. We thank Connor Fraser for research assistance. All errors are ours. Dang (daviddang@hotmail.com) and Li (fli@uwo.ivey.ca) are at Western University; Yang (cyang244@uwo.ca) is at Wuhan University.

## I. Introduction

Studies on the consequences and correlates of firm size can be traced back to a seminal article, Coase (1937), which raises the questions of how firm boundaries affect the allocation of resources and what determines firm boundaries. Both questions have received much attention in theoretical studies in economics and finance (e.g., Williamson (1975, 1986), Klein, Crawford, and Alchian (1978), and Grossman and Hart (1986)). Empirical researchers in corporate finance also consider firm size an important and fundamental firm characteristic, and, in many situations, observe a "size effect" – firm size affects the empirical results. For example, in capital structure, Frank and Goyal (2003) show that pecking order is only found in large firms; Rajan and Zingales (1995) discover that leverage increases with firm size. In mergers and acquisitions, Moeller, Schlingemann, and Stout (2005) find that target returns are higher for cash offers than for stock offers; Vijh and Yang (2013) document that for cash offers, targetiveness (the probability of being targeted) decreases with firm size, but for stock offers they find an inverted-U relation.

Although firm size matters in empirical corporate finance, the existing literature is silent on the rationale for using a certain measure of firm size, and no paper provides a comprehensive assessment of the sensitivity of empirical results in corporate finance to different measures of firm size. An exception is Vijh and Yang (2013), who provide a list of firm size proxies and their corresponding coefficients in the literature on takeover likelihood models. Their study indicates that the sign and significance of the coefficients on firm size in different papers are sensitive to which firm size measure is being used. While Vijh and Yang (2013) suggest that firm size measures should receive more attention, they do not compare the results based on the same regression or conduct a broader assessment of firm size measures in the corporate finance literature.

We use 20 representative specifications, in the areas of executive compensation, board of directors, corporate control, financial policy, payout policy, investment policy, diversification, and firm performance, to study the influences (sign sensitivity, significance sensitivity, and R-squared sensitivity) of employing different measures of firm size. For each specification, we employ natural logarithm forms of three firm size measures: total assets, total sales, and market value of equity. We choose these three measures because, according to our survey of 100 research papers, they are the most popular firm size proxies in corporate finance. However, other measures, such as number of employees and net assets, also appear in empirical work.

We choose the 20 representative specifications from Coles, Daniel, and Naveen (2006), Comment and Schwert (1995), Core and Guay (1999), DeAngelo, DeAngelo, and Stulz (2006), Graham, Li, and Qiu (2012), Harford (1999), Harford, Mansi, and Maxwell (2008), Lemmon, Roberts, and Zender (2008), Linck, Netter, and Yang (2008) and Mehran (1995). For brevity and data availability, we select the same papers as those in Coles and Li (2012). Coles and Li (2012) assess firm, manager, and time fixed effects in these 20 prominent areas in empirical corporate finance. On the one hand, our project is modest. Using our data sample with year fixed effects and industry fixed effects, our empirical models resemble the corresponding benchmark specifications in these papers. This allows an even-handed comparison between our results and those in the original papers and between results based on different firm size measures. On the other hand, our research thrust is ambitious in that we collect the data and perform the analysis for a large number of regression specifications across a wide spectrum of subfields in corporate finance.

Although all firm size measures are significantly correlated, they are theoretically *and* empirically different. The correlation coefficients range from 0.64 to 0.81 in our sample. Because

size is a firm fundamental variable, any subtle differences may have a critical impact on the dependent variable and other independent variables in the empirical study.<sup>1</sup> Our results indeed confirm this. The coefficients on regressors often change sign and significance when we use different firm size measures. We observe sign changes and significance changes in almost all areas except dividend policy and executive compensation. Unfortunately, this suggests that when using different firm size proxies, some previous studies are not robust.<sup>2</sup> Researchers should either use all the important proxies as robustness checks, or provide a rationale for using a specific proxy. Results based on a single size measure should be interpreted with caution. Second, the goodness of fit measured by R-squared varies significantly with different firm size measures. Some size measures appear more relevant than others in different research areas, implying that they are better control variables to reduce omitted variable bias and improve the estimation of the main coefficients of interest. And thus have different implications. The choice of these firm size measures can be a theoretical and empirical question. Finally, we use our results to provide guidelines on the choice of the size measure. The sensitivity of empirical results to different size measures not only provides guidance for researchers who must use firm size proxies in empirical corporate finance research, but also sheds light on future research that might incorporate measurement effects into other research fields, such as empirical asset pricing and empirical accounting.

---

<sup>1</sup> Per our results, the firm size measures are consistently one of the most significant independent variables in all subfields of corporate finance. In 18 out of 20 subfields, the size proxy is statistically significant at the 1% level.

<sup>2</sup> To provide even-handed comparisons, we attempted to use the same methodology and variable definitions in our experiments; we also tried the subsamples in the same time periods as in the original papers. The results are not qualitatively different, giving us confidence that our data and estimation are not so different from those papers. More importantly, we are not trying to argue against the results in the original papers. Instead, we test the sensitivity and robustness of the size measures in our larger, more comprehensive, and more recent data to raise awareness.

A few caveats should be noted. First, we do not employ all possible measures of firm size; we only study the three most popular measures. Researchers can use some alternative size proxies such as enterprise value (market capitalization plus net debt), the number of employees, total profits, or net assets (total assets minus total liabilities) when the main measures are not available or irrelevant (e.g., market cap for private firms and total sales for start-up firms). Second, we might omit some important representative papers in specific sub-fields due to data and time constraints. Third, some linear models may lose power if the true relation between firm size and the dependent variable is non-linear. Fourth, most of our empirical results are based on year fixed effects and/or industry fixed effects. Introducing other considerations, such as firm fixed effects (for consideration of within-firm variations rather than cross-sectional variations) or manager fixed effects (for emphasis on corporate governance issues such as managerial compensation), might change our results and have different implications.

The outline of the article is as follows. Part II includes our research motivation, a literature review, and the measures of firm size. Part III describes our data and the sample. Part IV provides a discussion of the empirical results. Part V concludes.

## **II. Framework for Analysis and Literature Review**

Coase (1937) states that firms are formed with boundaries to substitute markets in order to save transaction costs such as contracting and monitoring fees. For the effects of firm boundaries on firm behavior, Williamson (1975, 1986), Klein, Crawford, and Alchian (1978), and Grossman and Hart (1986) provide theoretical insights, while some recent works such as Holmstrom and Kaplan (2001), Robinson (2008), and Seru (2014) present empirical evidence that links the theory of firm and corporate finance to firm activities such as capital allocation.

Specifically, Bolton and Scharfstein (1998) review the relationship between corporate finance and the theory of firm and organizations.

As for the determinants of firm size, Kumar, Rajan, and Zingales (1999) comprehensively review the literature and classify the theories into four categories: technological theories (Lucas (1978), Rosen (1982), Kremer (1993), etc.), organizational theories (Williamson (1975, 1986), Klein, Crawford and Alchian (1978), Grossman and Hart (1986), Rajan and Zingales (1998b, 2001), Holmstrom (1999), Holmstrom and Roberts (1998), etc.), regulatory theories (Ringleb and Wiggins (1990), Hopenhayn (1992), etc.), and financial theories (Rajan and Zingales (1998a), etc.). Kumar, Rajan, and Zingales (1999) provide empirical evidence that the utility sector, R&D intensive industries, capital intensive industries, high wage industries, and industries that need little external financing all feature large firms.<sup>3</sup>

Several papers also investigate whether the measures of firm size are interchangeable in microeconomics and industrial organization, and these works are more associated with our goal to evaluate the effects of employing different firm size measures in empirical research. Smyth, Boyes, and Peseau (1975) first demonstrate that the measures of firm size are only interchangeable when more rigorous technical conditions than correlation are met. Smyth, Boyes, and Peseau (1975) show that economies of scale are sensitive to different firm size measures. Jackson and Dunlevy (1982) employ an asymptotically valid procedure to test the null hypothesis of orthogonal least squares suggested by Smyth, Boyes, and Peseau (1975). However, most financial studies usually use firm size measures without examining correlations and other interrelationships among different firm size measures. The empirical results in this paper support that the measures of firm size are indeed not interchangeable.

---

<sup>3</sup> Such evidence also motivates us to use industry fixed effects in our empirical investigations.

From the review above, we find that the overall assessment of firm size measures in empirical corporate finance has largely been ignored in the existing literature. However, this topic deserves attention. In most prominent areas of empirical corporate finance research, finance scholars employ firm size as an important firm characteristic, and in many situations, finance scholars have observed the "firm size matters" in determining the dependent variables. For example, it is well recognized that top-management compensation level increases with firm size (Jensen and Murphy (1990), Core, Holthausen, and Larcker (1999), etc.). Baker and Hall (2004) find that CEO marginal products increase substantially with firm size. Gabaix and Landier (2008) and Gabaix, Landier, and Sauvagnat (2014) show that small differences in CEO talent can result in substantial differences in CEO pay through the effect of firm size; in particular, larger firms usually have more skilled managers (Himmelberg and Hubbard (2000)).

Although the majority of the literature takes for granted that the choice of firm size measures is not a vital concern, we doubt the existence of selection bias of empirical results in some papers. Recent works (e.g. Vijh and Yang (2013, Appendix 2)) find that the sign and significance of the coefficients of size proxies in the literature of mergers and acquisitions are sensitive to different firm size measures. While Vijh and Yang (2013) indicate that firm size measures should receive more attention, they are silent on the assessment of firm size measures based on the same regression and the comprehensive assessment in broad corporate finance literature. In addition, Vijh and Yang (2013) have little to say on the sensitivity of the coefficients of regressors other than firm size when different firm size measures are employed. These limitations in the existing literature motivate us to investigate the effects of different size measures comprehensively.



For the purpose of conducting a comprehensive empirical assessment of firm size measures in different sub-fields of empirical corporate finance, we follow the methodology of Coles and Li (2012), covering 20 prominent research areas in corporate finance: financial policy (book leverage, market leverage, and cash holdings), payout policy (dividend dummy), investment policy (CAPEX, R&D, and firm risk), diversification (Herfindahl index and business concentration), profitability (ROA, which is the sum of market capitalization of equity plus liabilities divided by total book assets, and ROA, which is the ratio of net income to total assets), mergers and acquisitions and corporate control (bidder, target, and poison pills), managerial compensation and incentives (delta, vega, and pay level), and board of directors (board size, board independence, and CEO duality).

We employ three firm size measures: total assets, total sales, and market value of equity. According to our survey, in which we investigate 100 empirical papers from top finance, accounting, and economics journals that use firm size measures on the topics of empirical corporate finance, these three measures are the most popular firm size proxies in empirical corporate finance research. We collect a total of 100 papers through *Google Scholar* by searching subfield key words, and the results are listed by descending number of citations. We only choose the papers that appear in top journals and use firm size measures in empirical studies. The papers are distributed across extensive areas in corporate finance, including capital structure, debt policy, payout policy, cash holdings, corporate investment and financial constraints, cross listings, CEO turnover, CEO compensation, board of directors, law and finance, ownership structure, mergers and acquisitions, and corporate control (see Appendix for detailed information on these papers). We find that these three firm size proxies are used in 85 papers out of the 87 papers that use single measures, and the remaining 13 papers use multiple measures for

robustness checks. Among these 87 papers, 49 papers use total assets, 20 papers use market capitalization, 16 papers use sales, and 2 papers use number of employees. We are aware that other measures, such as number of employees and net assets, also appear, though infrequently, in empirical finance works, but for conciseness we only use these three measures. In addition, most papers in empirical corporate finance use the logarithm form of firm size measures. In the 100 papers we surveyed, only 3 papers use the original form of the three size measures. This suggests that it is a rule of thumb in corporate finance to use log form to mitigate the high skewness of firm size data.

Interestingly, *Forbes Global 2000* uses four measures (assets, sales, profits, and market cap) to rank all the large companies in the world, and *Fortune 500* uses just two measures (sales and profits). Both employ sales and profits, but profits seldom appear as a proxy for firm size in academic research.

Every firm size measure exhibits advantages and disadvantages, and no measure can capture all characteristics of "firm size". Generally, total assets measures total firm resources, market capitalization involves firm growth opportunities and equity market condition, and total sales measures product market competition and is not forward looking. In addition, researchers can use the number of employees, total profits, and net assets when the main measures are not available or irrelevant (e.g., market cap for private firms and total sales for start-up firms). Moreover, Hart and Oulton (1996) argue that net assets can be negative but sales are always positive. They also point out that number of employees does not include the number of part-time workers, but nowadays part-time workers play an important role. Because every measure has pros and cons, Hart and Oulton (1996) suggest that, in practice, choosing which measure to use depends on data availability. In addition, we think the choice of firm size measures also depends

on the purpose of the specific study. For example, Prowse (1992) applies different firm size measures as the research purpose changes from the ownership of equity to the ownership of asset.

In sum, we find that the existing literature has little to say about the rationale of using a certain measure of firm size for specific corporate finance research, and no paper provides a comprehensive assessment of the sensitivity of empirical results in corporate finance to different measures of firm size. This hole in the literature motivates us to find evidence for a  $\delta$   $\circ$   $g$   $c$   $u$   $w$   $t$   $g$   $\circ$   $g$   $p$   $v$   $"$   $g$   $h$   $h$   $g$   $e$   $v$   $\ddot{o}$   $"$   $k$   $p$   $t$   $h$   $e$   $\delta$   $u$   $k$   $|$   $g$   $"$   $g$   $h$   $h$   $g$   $e$   $v$   $\ddot{o}$ , and to provide a general guideline to researchers who must use firm size, as a key variable or control variable, in their empirical corporate finance studies.

### **III. The Data**

We extract data from multiple sources. Corporate governance data are from RiskMetrics Governance, director data are from RiskMetrics Directors, stock daily returns and prices are from CRSP, company diversification data are from Compustat Segment, corporate bond data are from Compustat Ratings, institutional holdings data are from Thomson Reuters, Executive data  $\hat{o}$  up to five top executives per firm  $\hat{o}$  are from ExecuComp, M&A deals and corporate control data are from SDC, and all other financial items are from Compustat Fundamentals. We restrict the observations to only those that match North American data from CRSP and Compustat for firms with fiscal years 1993-2006. In line with conventional tradition, we exclude data from the financial and utility sectors. See Table 1 for summary statistics of all the variables featured in our representative specifications from corresponding benchmark papers. Specifically, we report data properties and bivariate scattergrams of the alternative firm size measures in logarithm numbers for the regressions of firm  $rg$   $h$   $q$   $t$   $\circ$   $c$   $p$   $e$   $g$   $"$   $*$   $V$   $q$   $d$   $k$   $p$   $\hat{o}$   $"$   $S$   $"$   $c$   $p$   $f$   $"$   $T$   $Q$   $C$   $+$  as an example.

Table 2 Panel A reports summary statistics of firm size measures for both raw numbers and logarithm numbers. Panel B presents the Pearson correlation coefficients of firm size measures across raw numbers and logarithm numbers. Figure 1 shows bivariate scattergrams of alternative firm sizes measured in logarithm numbers, which we employ in the regressions. We find that the correlation coefficients among log (assets), log (sales), and log (market value of equity) are between 0.77 and 0.92, and those among raw numbers are between 0.64 and 0.81. The highest correlation coefficient is between log (assets) and log (sales) (0.92), and the lowest correlation coefficient is between sales and log (market value of equity) (0.50). These correlations indicate that although all the size measures are significantly correlated, they are different and some are more correlated than others.

We also show the trends of the three different firm size measures over our testing period in Figure 2. Figure 2A is expressed in logarithm form and Figure 2B in original form in 2006 dollars. The average market capitalization in 2002 decreased dramatically, consistent with the dot-com bubble burst. The bottom line is that time trends appear different for different measures,

#### **IV. Methodology and Empirical Results**

We adopt the empirical methodologies in the benchmark papers by employing conventional short-panel techniques for basic empirical analysis. For each specification, we apply both basic OLS regressions and industry fixed effect regressions. Time fixed effects are included in every regression. We use industry fixed effect because some benchmark papers employ 2-digit SIC controls (e.g. Coles, Daniel, and Naveen (2006)) and others only include industrial firms (e.g. DeAngelo, DeAngelo, and Stulz (2006)) or manufacturing firms (e.g.

































































































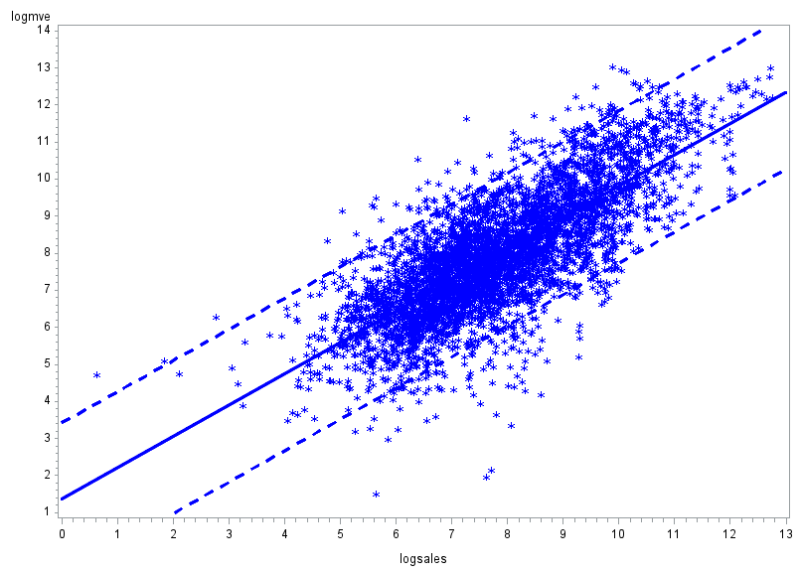
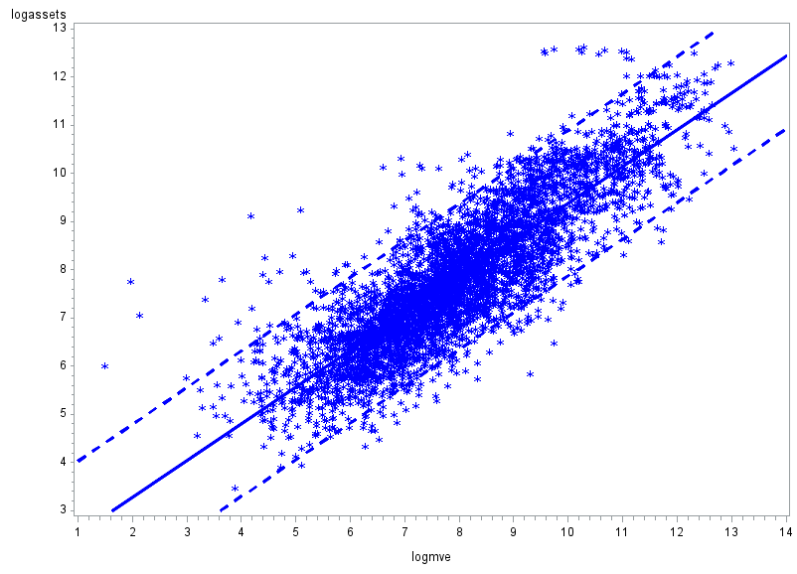








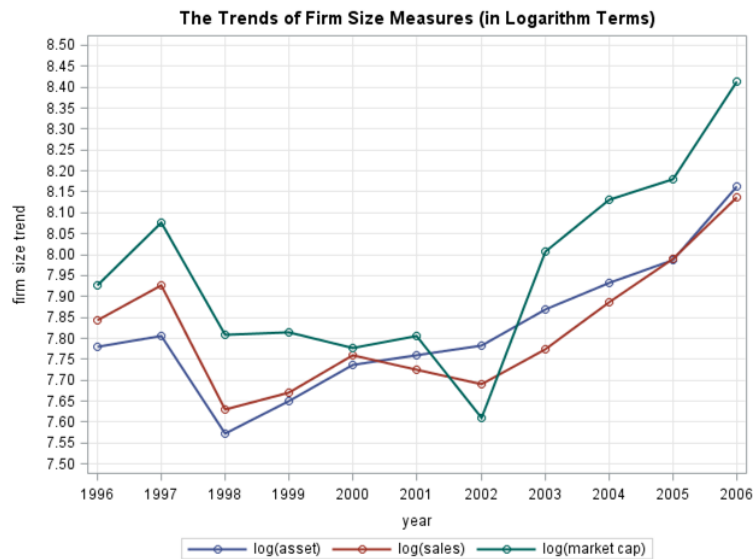




**Figure 2: Time Series of Alternative Firm Size Measures**

Figure 2 provides the time series of the average firm size measures for all the firms in the data sample for firm performance. Figure 2A shows trends in logarithm form, and Figure 2B shows trends in original form (in 2006 dollars).

**Figure 2A: The Trends of Firm Size Measures (in Logarithm Terms)**



**Figure 2B: The Trends of Firm Size Measures (in Original Terms)**

