THE CORRELATION BETWEEN FIRM PROFITABILITY AND FIRM SIZE

Charles Kile, Western Kentucky University
Jonathan Ross, Western Kentucky University

ABSTRACT

We document a statistically negative correlation between firm profitability (measured as the return on assets ratio (ROA)) and firm size (the denominator of ROA) over the 1966-1974 time period. This negative correlation ranges from -0.3 to -0.05. The correlation switched to statistically positive beginning in 1975, and has experienced an increasing trend through 2022, where the correlation was 0.55. When examining the extreme deciles of the U.S. publicly traded market, we find that the firms in the extreme small(large) size decile have experienced statistically positive(negative) correlations between ROA and size. This suggests that small firms have experienced economies of scale while large firms have experienced diseconomies of scale. Firms in the middle size decile have correlations much closer to zero, on average, over the 1966-2022 sample time period. We find that the trend in correlation within industry mirrors the overall trend. We showt hat the correlation between ROA and size, within industry, is positively related to industry concentration, controlling for the number of firms in the industry.

INTRODUCTION

Our study is motivated by the fact that return on assets (ROA) is one of the most commonly used measures of profitability by which firms are compared. The standard assumption is that dividing earnings by assets enables an apples-to-apples comparison of earnings per dollar of assets between two or more firms. The implicit assumption that underpins this computation is that ROA and its denominator, size, are not correlated with each other. That is, the difference between the ROA's of two firms is not correlated with the difference between the size of the two firms. We show that this assumption does not hold for U.S. publicly traded firms.

Further motivation for our study lies in the general fact that accounting educators teach their students ratio analysis for the purpose of comparing firms with themselves over time (horizontal analysis) and for the purpose of comparing across firms holding time constant. The underlying assumption with ratio analysis is that the ratio is not correlated with its denominator. Students then go on to become financial analysts and/or researchers and carry this assumption with them. Evidence of this lies in the fact that financial analysts still do relative comparison of companies using financial ratios without further adjustment to said ratios, assuming that these ratios admit unbiased relative comparison. Reseachers also demonstrate their adherence to the underlying assumption by frequently adding ratios (scaling usually by size) to their regression models as control variables, without also adding the denominator. This model misspecification can lead to correlated omitted variable bias in estimating the coefficients on the variables of interest in the regression model.

Our paper shows that one of the most commonly used ratios in ratio analysis, ROA, is significantly correlated with its denominator. Specifically, we find that there has been a structural shift in the correlation between ROA and firm size over our 1966-2022 sample time period. The correlation between ROA and firm size averaged a significant -0.162 over 1966-1974 and a significant 0.401 over 1975-2022. In the past 25 years, the correlation between ROA and firm size has been hovering in the range of [0.5,0.6]! When breaking the sample into deciles on firm size each year, we find that the largest firms (decile 10) experience a statistically negative correlation (mean of -0.125) over the sample time period while the smallest firms (decile 1) experience a statistically positive correlation (mean of 0.215) over the sample time period. ROA and size is statistically positively correlated (mean of 0.028) for the medium-sized firms (decile 5) over the sample time period. Thus there is a non-linear relationship between ROA and firm size. The positive sensitivity of ROA to firm size in small-medium firms is greater than the negative sensitivity of ROA to firm size in large firms. This provides evidence consistent with the theory of Alchian (1965) and Williamson (1963). Namely, there is a greater separation between management and ownership with large firms and thus managerial utility maximization is replacing profit maximization as the objective function of the firm. Thus, as firm size increases, profit decreases and we observe a negative correlation between ROA and firm size, because managers are maximizing a different objective function. Furthermore, we provide evidence consistent with the assertion in Amato and Wilder (1985) that there may be a threshold size level, above which growth in size leads to decreasing return on assets. This threshold size level in our study is relative to the year and occurs at approximately the 71st percentile of total assets in a given year. Size increases above this level lead to negative correlation while size decreases below this level lead to positive correlation between ROA and size.

Our study is the first, to our knowledge, to document a negative correlation between ROA and size in the distant past and a positive correlation between ROA and size in the past ~50 years in the cross-section of U.S. firms. Furthermore, our study is the first, to our knowledge, to document the interesting differences in the correlation trend between the extreme small and large decile firms. We shed light on the notion that the smallest firms experience increasing returns to scale while the largest firms experience decreasing returns to scale. We find evidence of the approximate inflection point (size threshold), above which further increases in size lead to decreasing operating income.

We also examine whether the overall trend in correlation between ROA and firm size holds within industry. Specifically, we test whether the same patterns discussed above hold when we constrain the correlations to be measured only between firms in the same industry. We find the same patterns hold almost exactly as when we don't constrain the correlations to be measured between only those firms which share the same industry. We do find evidence that there are some industries whose firms persistently display a negative correlation between ROA and firm size and others whose firms persistently display a positive correlation. Finally, we establish, within a regression framework, that the correlation between ROA and firm size is strongly statistically positively associated with the concentration of the industry and the number of firms in the industry. This finding is consistent with and is an incremental contribution to the findings of Grullon et al. (2019).

RELATED LITERATURE

Lev and Sunder (1979) point out (see page 188) that researchers and practitioners often implicitly assume a certain relationship between the numerator and denominator when using firm ratios in their analyses. Often, the numerator variable is a firm-specific accounting measure, such as earnings, and the denominator is a measure of firm size, such as total assets or equity. The objective for practitioners is usually to compare two firms on the basis of profitability, liquidity or solvency. The objective for researchers is usually to control for the well-established effect of firm size on most would-be dependent variables in a regression analysis setting, in order to isolate the effect of a particular explanatory variable. The implicit assumption is that there is a constant linear relationship between the numerator variable, y and denominator, x. Equation (1) is, therefore, assumed to hold, with beta a constant.

$$y = \beta x \tag{1}$$

Equation (1) implies the ratio, y/x, is constant and therefore not correlated with x. Lev and Sunder (1979) (see section 2.1 on page 190) urge practitioners and researchers to carefully consider whether there is a theoretical relationship between the numerator and denominator before indiscriminately using ratios in financial analyses.

Hall and Weiss (1967) test the economies of scale hypothesis put forth in Baumol (1967). Namely, that "large firms have all of the options of small firms, and, in addition, they can invest in lines requiring such scale that small firms are excluded" (see 2nd paragraph on page 319). They find evidence in favor of the economies of scale hypothesis. Specifically, they find that return on assets increases with size (measured as the natural logarithm of total assets), after controlling for industry concentration and other variables. None of their control variables, by the way, were statistically significant. Their sample, however, was very small and only included 326 of the Fortune 500 firms over the short sample period 1956-1962. Thus, in the largest size echelon, over a short time period, return on assets and size were statistically positively correlated.

Amato and Wilder (1985) use a sample that covers a wide range of firm sizes of U.S. manufacturing firms, over the time period 1966-1975. They posit that "the relationship between firm size and profit rates may be positive over some firm size ranges and negative for others so that the relationship could be non-linear". Smaller firms are able to realize economies of scale while larger firms experience diminishing marginal returns (see page 183). Thus Amato and Wilder (1985) suggest there is a threshold size level, above which growth in size leads to decreasing return on assets. Reasoning provided in Alchian (1965) (see pages 35-36), which is based on the analytical model of Williamson (1963), suggests that managerial utility maximization may replace profit maximization as the firm's objective function. This occurs because of a greater degree of separation between ownership and management brought about by increases in firm size. Thus, as firm size increases, profit decreases because managers are maximizing a different objective function. Amato and Wilder (1985), however, find no statistical relation between firm size and profitability.[1]

A classification system of theories regarding why firms exist and the determinants of firm size was put forth in Rajan et al. (2001). They classify these theories into three categories: technological, organizational and institutional. Each category has different implications for the relation between firm size and profitability. Overall, the pre-dominant theory that has emerged has been the following. A small firm will experience increases in profitability as firm size increases, due to economies of scale. At some optimal firm size, the combined effect of much higher organizational costs outweigh the economies of scale that can be realized and further increases in size lead to a decrease in profitability. (Etebari et al. [2010])

There have been a number of more recent empirical studies which have studied the correlation between firm profitability and firm size. A large number of these have been done by researchers abroad using foreign (to the U.S.) samples and the results are mixed. Using a sample time period of 1987-2002, Becker-Blease et al. (2010) examine the relation between firm profitability and firm size for U.S. firms within the same SIC four-digit manufacturing industry. A negative(no)(positive) relation between firm profitability and size is found in 47(52)(11) of their industries which indicates that the relation is industry-specific. This result is in line with what we found in our much larger sample, across all SIC two-digit industries (not constrained to only manufacturing). Thus, our results extend their results to a wider population.

Using a sample of 200 companies active on the Istanbul Stock Exchange over 2008-2011, Dogan (2013) finds a statistically positive correlation between firm ROA and and total assets of 0.16. The median total assets of firms in his sample is $19.57 million which, compared to our sample, puts his entire sample of companies in our first decile (whose cutoff was < $22.60 million). In comparison, the correlation between ROA and total assets of our decile 1 firms was 0.23. Other studies that have found a positive relationship between firm profitability and firm size in foreign firms are: Babalola (2013) -- Nigeria, Isik et al. (2017) -- Turkey and Rahman and Yilun (2021) -- China. Two studies have found no significant correlation between firm profitability and size in Sri Lankan firms: Niresh and Thirunavukkarasu (2014) and Abeyrathna and Priyadarshana (2019).

A recently published article that serves as a good example of how one should deal with ratios in a regression is Seissian (2024) who examines the determinants of internet financial reporting (IFR) using a sample of companies which have lagged the rest of the world in doing so; namely public companies which orginated in the Middle East and Gulf Cooperation Council (GCC). In a multiple linear regression framework, she finds, among other variables, that profitability (return on assets ratio) negatively impacts IFR while size positively impacts IFR. The fact that she includes both variables in her regression is good because if she had omitted either of them the coefficient on the remaining variable would have been biased. Another study, Susetyo (2023) examines the effect of both firm profitability and firm size on leverage (debt-to-equity ratio). Both explanatory variables are included in the regression together and collectively impact the dependent variable.

Finally, motivating our within-industry analysis, Grullon et al. (2019) examine how the concentration of industries has changed in the U.S. over the time period 1972-2014. They show

[1] Other older studies which empirically examine the relationship between firm profitability and firm size are Ravenscraft (1983), Smyth et al. (1975), Stekler (1964) and Steindl (1945).

that industry concentration has increased substantially while many firms have dissapeared from U.S. product markets and there has been a large-scale consolidation of firms. The large-scale consolidation has led to a three-fold increase in the median firm's size over the period 1994-2014. They examine whether this increase in concentration is related to changes in firm profitability and find a positive relationship between industry concentration and firm profitability.

HYPOTHESES

Our study is similar in nature to the previously discussed studies in that we examine the empirical relationship between firm size and firm profitability. We expand the prior literature by covering a much larger time period of 1966-2022. We first examine the logical condition put forth in Lev and Sunder (1979) regarding the empirical relationship between the numerator of ROA and the denominator. Specifically, ROA and firm size will be correlated if operating income is not equal to a constant multiple of firm size. Our first hypothesis is therefore given below.

H1: The association between firm operating income and firm size is not constant over time.

Our next hypothesis is the same as Hall and Weiss (1967), but we test it over a much larger sample time period and with firms of all sizes (instead of only the largest firms). Amato and Wilder (1985) also test the following hypothesis (but stated differently) and don't find statistical evidence in favor of it. We challenge their finding with our second hypothesis below.

H2: The difference between two firms' respective ROAs is correlated with the difference between their respective sizes.

We state H2 in the context of comparing two firm's ROAs to emphasize the importance of how ROA is often used in practice. If we find evidence that the average pair of firm's difference in ROA is correlated with their respective difference in size, then ROA is correlated with firm size, and using ROA as a basis for comparison, without further adjustment, is called into question. We don't specify a direction for the correlation hypothesized in H2. In the spirit of Amato and Wilder (1985), we specify a non-directional hypothesis. We allow for the possibility that, for some firms, ROA and size may be negatively correlated, while for others ROA and size may be positively correlated. If we find a positive(negative) correlation between ROA and size for small(large) firms then we will provide empirical evidence that is consistent with the theory put forth in Alchian (1965) and Williamson (1963).

Our final hypothesis is informed by the empirical findings of Grullon et al. (2019). Specifically, they find that industry concentration has substantially increased over the time period 1994-2014 and this mirrors the trend of firms becoming three times larger over this same time period. Since Grullon et al. (2019) find that firm size has greatly increased, industries have become more concentrated and this increased concentration has led to increases in firm profitability, holding other factors constant, the correlation between ROA and firm size should thus be positively related to industry concentration.

H3: The correlation between ROA and firm size is positively associated with industry concentration, controlling for industry size.

Industry concentration is mechanically, positively related to the number of firms in the industry (industry size). To the extent that the correlation between ROA and firm size is also related to industry size, we should control for industry size in our test of H3. Grullon et al. (2019) find a positive correlation between ROA and firm size while they are examining industry concentration. In H3, we are not predicting the sign of the correlation between ROA and firm size. Rather, we are simply predicting that whatever the correlation is between ROA and firm size (whether negative or positive), this correlation will increase with industry concentration, controlling for industry size.

METHODOLOGY

Our method for ascertaining the correlation between firm ROA and firm size directly follows our statement of H2. For each year of our 1966-2022 sample time period, we compute the difference in ROA for each possible pairs of firms in that year and the corresponding difference in size (measured as the natural logarithm of their average total assets). Next, we measure the Pearson (Spearman) correlation between the ROA difference and size difference for each year. The plot of this correlation over our sample time period is plotted in Figure 3.

For example, suppose that in 1966 there were four firms in our sample who had the appropriate data to compute ROA. Suppose those four hypothetical firms had the following data shown in Table 1. The data used in Table 1 is actual data pertaining to four of our sample firms.

Table 1
Hypothetical Example Used to Illustrate Our Methodology (step 1)

Firm	Year	Inc. ($mil)	Beg. TA ($mil)	End TA ($mil)	Avg. TA	ROA	ln(avg. TA)
1	1966	2.010	42.700	68.600	55.650	0.036119	24.742348
2	1966	2.832	29.672	33.326	31.499	0.089908	24.1732222
3	1966	39.860	507.700	530.900	519.300	0.076411	26.975748
4	1966	-4.182	67.200	51.500	59.350	-0.070463	24.806718

where “Inc. ($mil)” is the firm's net income after depreciation and amortization, in millions of dollars, “TA” is total assets, “ROA” is return on assets and “ln” represents the natural logarithm function. The reader can verify the average total assets, ROA and ln(avg. TA), from the other numbers given in Table 1.[2] We next compute the difference in ROA and the difference in Size between all possible pairs of firms in Table 2.

[2] All of the results in the paper are robust to alternative measures of income.

Table 2
Hypothetical Example Used to Illustrate Our Methodology (step 2)

Firm-Pair	Difference in ROA	Difference in ln(avg. TA)	Difference in Avg. TA ($mil)
1	1966	2.010	42.700
2	1966	2.832	29.672
3	1966	39.860	507.700
4	1966	-4.182	67.200

We measure Size in our paper as the natural logarithm of average total assets.[3] Finally, we compute the Pearson (Spearman rank) correlation between the "difference in ROA" and the "difference in ln(avg. TA)" and get 0.459 (0.143).[4] Every time we refer to the "correlation between ROA and Size" in our study, we are referring to the correlation that we computed using the method just described.

Of course, an easier method would be to just compute the correlation between "ROA" and "ln(avg. TA)" in Table 1. Doing so for our hypothetical example yields a Pearson (Spearman) correlation of 0.215 (-0.400).[5] We feel this is not as precise a method to compute the correlation between the two variables because it, in essence, aggregates and thereby loses the information that the pair-wise differences in ROA and Size capture.

For our industry analysis, we used the same method as illustrated with our four-firm hypothetical example. The only difference being that we require the firms to be in the same industry before computing their pair-wise difference in ROA and Size, respectively.

TESTS OF HYPOTHESES AND RESULTS

Tests of H1

Our sample selection begins with all firms in the Annual Fundamentals file of Compustat for which there are non-missing observations on the variables: operating income (oiadp), total assets (at) and the 4-digit standard industrial classification code (sic). Our sample time period is 1966-2022. We start with 1966 because years before that have drastically reduced Compustat coverage. We delete firm-years with non-positive total assets and firm-years with zero operating income. We compute the return on assets (ROA) in a given firm-year as the operating income divided by the average assets. We next delete firm-years whose absolute value of ROA is greater

[3] All of our results, however, are robust to measuring Size as just the average of total assets.

[4] Note, that if we instead compute the correlation between the "difference in ROA" and the "difference in avg. TA", we get 0.605 (0.143). We tried this as an alternative method in our paper and the same pattern emerged in Figure 3 but was just shifted upward. The other results remained qualitatively the same.

[5] We did that as an alternative method by which to compute the "correlation between ROA and Size" and all of the results reported in the paper were qualitatively the same.

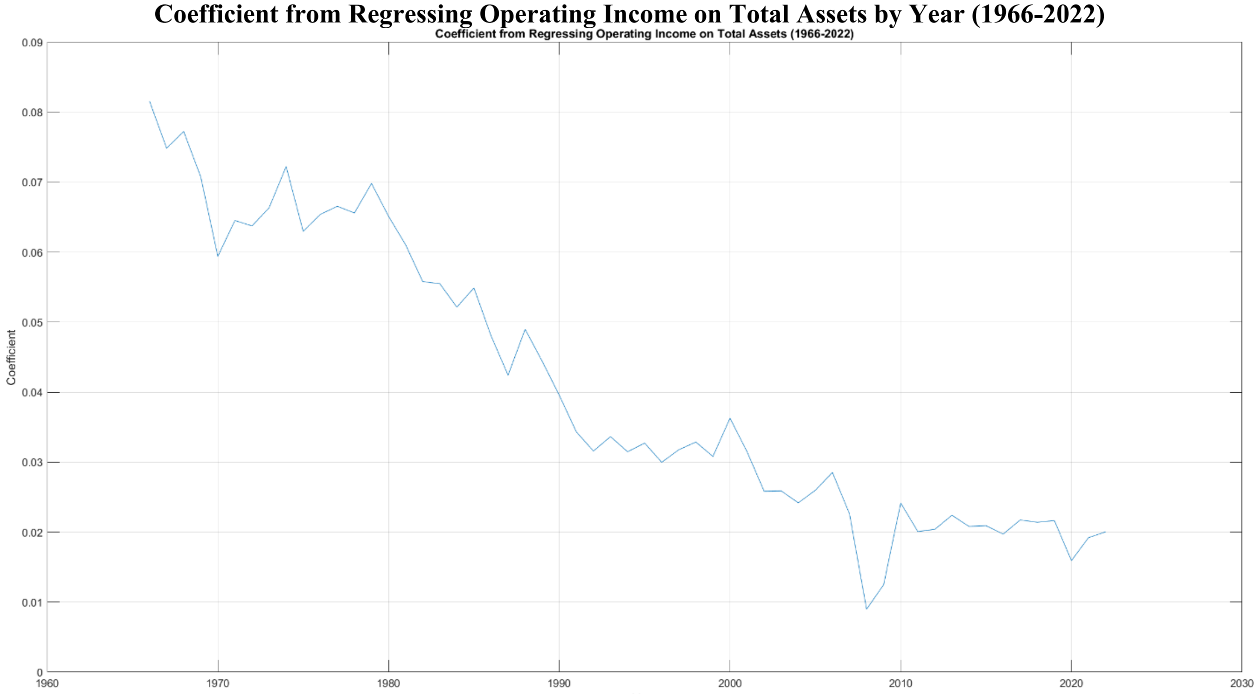
than 2.5. This constraint intends to remove extreme ROA outliers and reduces our sample by about 5000 firm-year observations (or around 1.4%). For our industry tests, we further constrain our sample such that all firm-years without at least 5 observations in the same SIC 2-digit industry are removed. This constraint reduces the sample by an additional ≈ 800 firm-year observtaions. Our final sample, on which we conduct all of our tests, consists of 357,037 firm-year observations over the 57-year time period 1966-2022.[6]

To test H1, we estimate regression equation (2),

$$OpInc_i = \beta_1 Assets_i + \varepsilon_i \quad (2)$$

for each year of our sample time period, where $OpInc_i$ and $Assets_i$ are the operating income and total assets for firm i. Notice that we omit the constant in estimating regression equation (2) since we are testing the constraint of Lev and Sunder (1979).[7] Figure 1 plots β_1 for each year of our sample time period.

Figure 1
Coefficient from Regressing Operating Income on Total Assets by Year (1966-2022)



Notice how β_1 is not constant over the sample time period but starts at around 0.08 in 1966 and has decreased steadily and leveled off at around 0.02 in 2022. Thus each $1 of assets

[6] All of our results are robust to using other measures of firm size, such as net sales revenue and total shareholder's equity and other measures of return, such as net income.

[7] See the discussion in Section 2.1 on page 190.

generates about four times less dollars of operating income than it used to. The coefficient was strongly, statistically greater than zero in every year. The mean(median) t-stat of β_1, over the years, is 85.37(86.53). A t-test on the time series of β_1, from estimating regression equation (2), yields a t-stat of 15.35. The mean(median) R^2 from estimating regression equation (2) is 52.18%(53.33%). Figure 1 provides strong evidence that the criteria set forth in Lev and Sunder (1979) does not hold in U.S. firms. Therefore, we provide evidence in favor of H1 and the relationship between total assets and operating income is not constant.

Tests of H2

The evidence found to support H1, by itself, implies that ROA will be correlated with total assets (size) and thus provides evidence in favor of H2. To directly test H2, we compute the difference in ROA for each pair of firms and the respective difference in size (measured as the natural logarithm of total assets) for that same pair of firms. We then compute the Pearson and Spearman correlations between these pair-wise differences in ROA and corresponding pair-wise differences in size in each year and plot the results in Figure 3. Figure 2 plots the number of sample firms in the market, each year, over our sample time period.

Figure 2
Number of Sample Firms by Year (1966-2022)

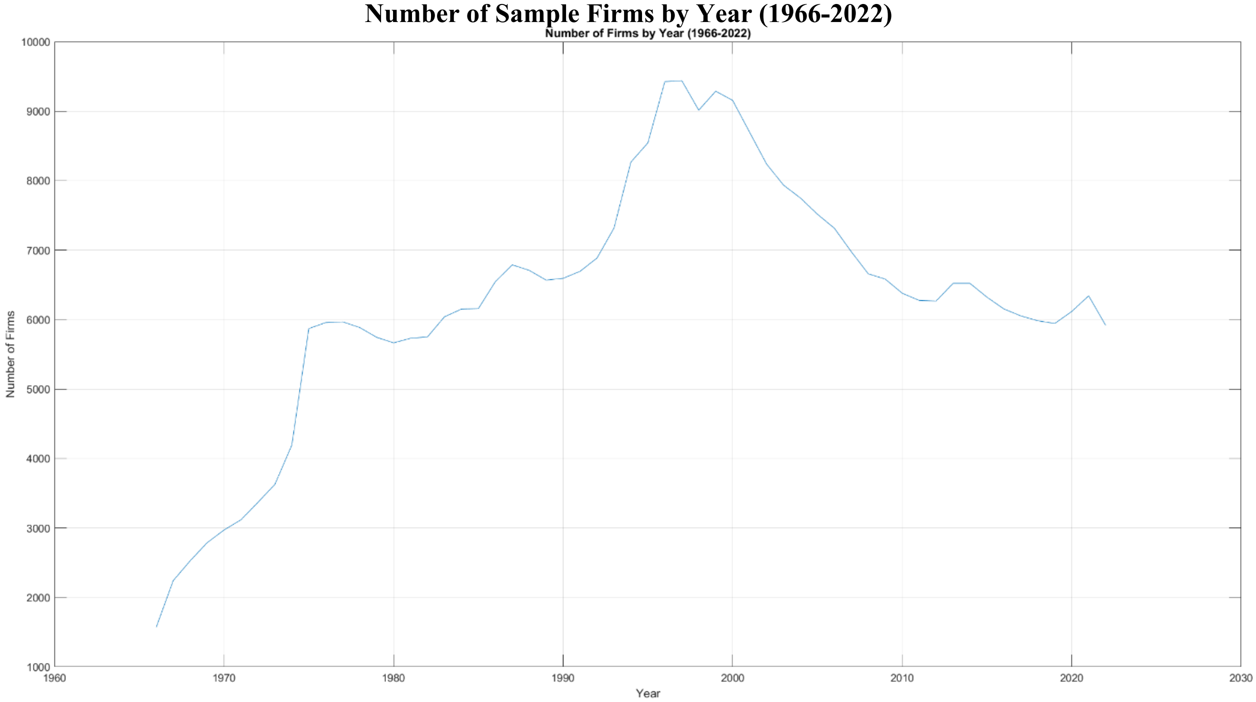
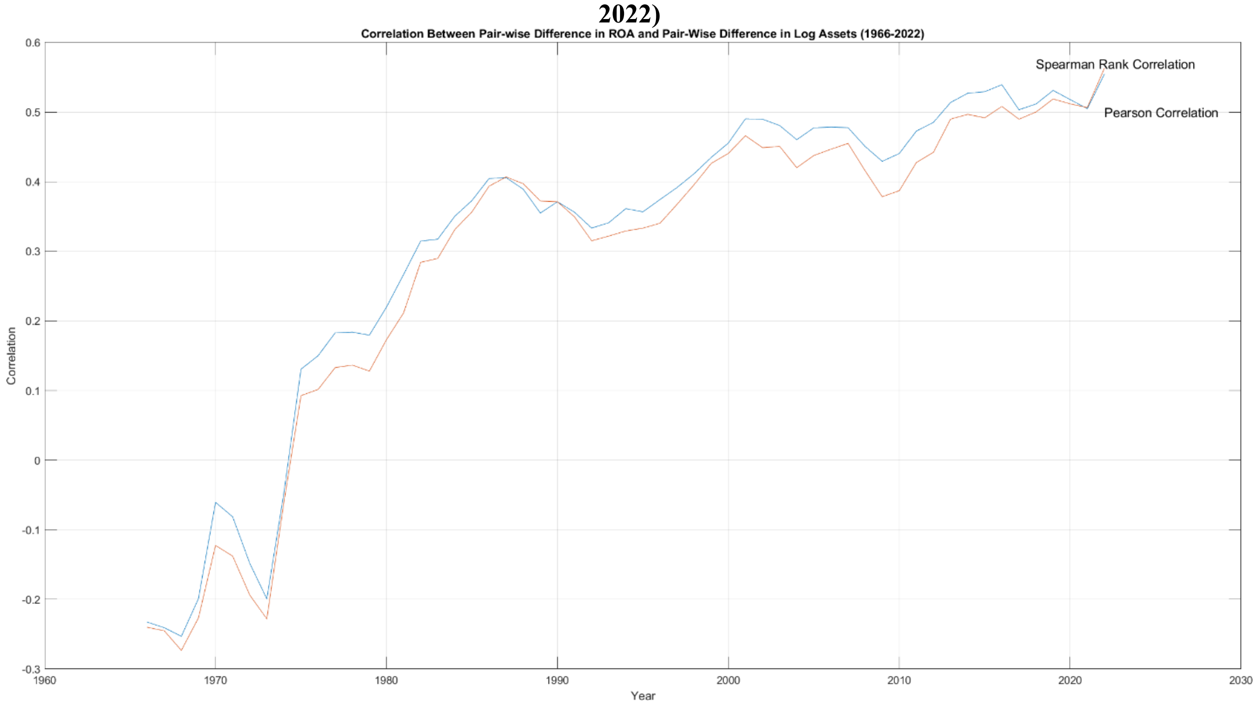


Figure 3
Correlation Between Pair-wise Difference in ROA and Pair-wise Difference in Log Assets by Year (1966-2022)



The plot in Figure 2 reveals the commonly-documented trend of publicly traded U.S. market firms. The size of the market peaked around 1997 and declined drastically after the dot com bubble burst. Figure 3reveals that the correlation between ROA and size was negative over 1966-1974 and has been positive and increasing ever since.[8] The mean(median) Pearson correlation value over the 1966-1974 time period is -0.162(-0.199), while the mean(median) correlation value over the 1975-2022 time period is 0.402(0.420). The mean(median) correlation value over the entire sample time period is 0.313(0.389). A t-test of the time-series of correlations plotted in Figure 3 yields a t-stat of 10.12. Therefore, we provide strong evidence in favor of H2 that the correlation between ROA and size is not zero. Our study is the first, to our knowledge, to document a negative correlation between ROA and size in the distant past and a positive correlation between ROA and size in the past ~50 years.

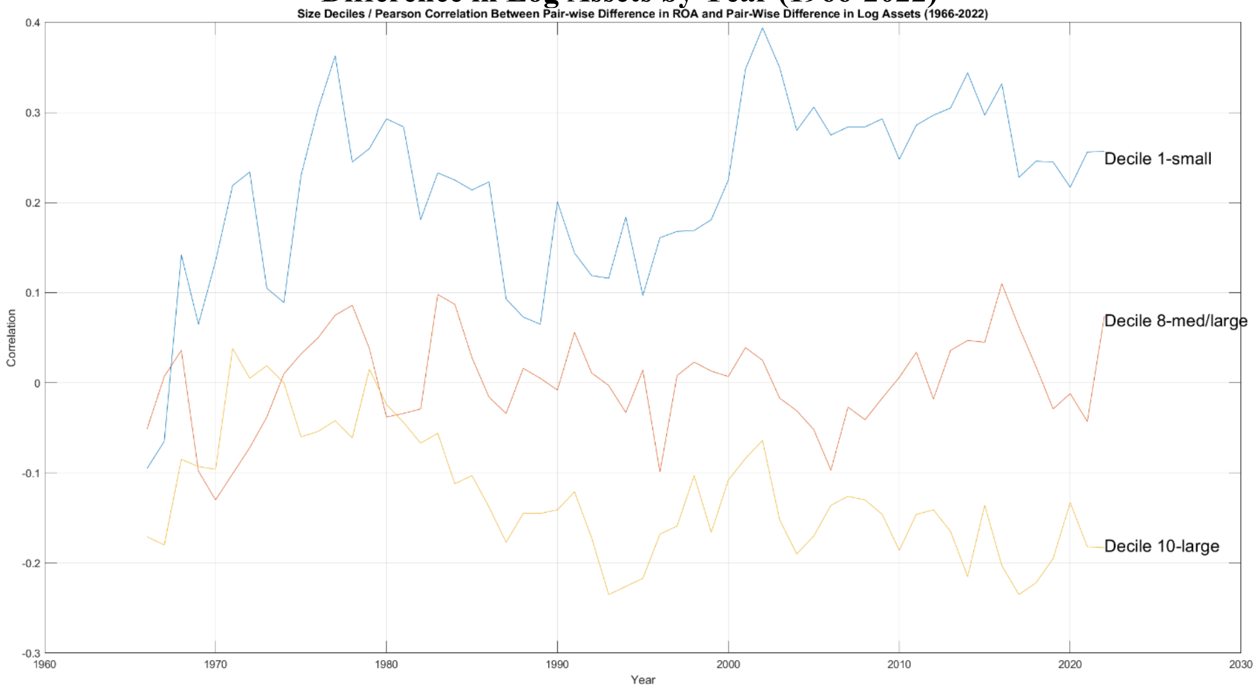
We further examine H2 by plotting the correlation between ROA and size by decile of size in Figure 4.[9]

---

[8] Whether you use the simple correlation of ROA with size in each year, or the correlation of the pair-wise differences in ROA with the corresponding pair-wise differences in size, you arrive at almost exactly the same plot.

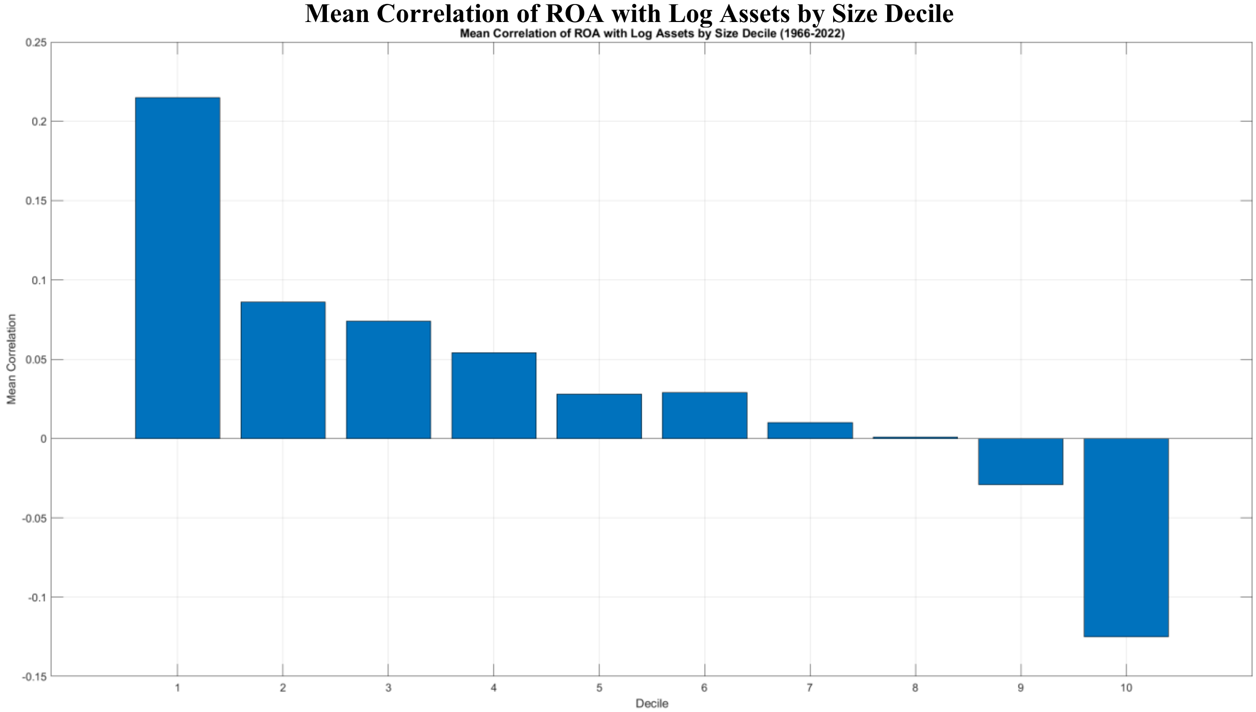
[9] Whether we measure size using raw total assets or the natural logarithm of total assets, we obtain qualitatively similar results.

Figure 4
Size Deciles of Pearson Correlation Between Pair-wise Difference in ROA and Pair-wise Difference in Log Assets by Year (1966-2022)



Notice the sign and the pattern of correlation changes, depending on the size decile. In the extreme smallest size decile, ROA and size are increasingly positively correlated with each other over time. In the extreme largest size decile, ROA and size are increasingly negatively correlated with each other over time. The correlation between ROA and size for firms in decile 8 has hovered around 0 over time. Figure 5 plots the mean correlation between ROA and size by size decile over our sample time period. Notice how there is a non-linear relationship between the correlation and the size decile. That is, as size increases, the correlation changes from positive to negative in an unequal manner as move from one size decile to the next. Even size deciles 5 and 6 still have statistically positive mean correlation between ROA and size.

Figure 5
Mean Correlation of ROA with Log Assets by Size Decile



The diagonal entries of Table 3 display the t-stats from one-sample t-tests that compare the mean of the time series of correlations of ROA with size to zero, for each size decile. Notice that the mean of the correlations between ROA and size, over our sample time period, is insignificantly different from zero in only size decile 7 and 8. The off-diagonal entries of Table 3 display the t-stats from paired-sample t-tests that compare the means of the time series of correlations of ROA with size between all possible pairs of size deciles. Notice that most of the t-stats are significantly different than zero. This indicates that the mean of the time series of correlations between ROA and size of one size decile is statistically different fom the mean of the time series of correlations between ROA and size of another size decile for almost all of the 45 possible pairs of size deciles, over our sample time period. The only exceptions are for the following pairs of size deciles: {(2,3), (3,4), (5,6), (5,7), (7,8)}.

Table 3
One and Two-Sample t-tests of Correlation Between ROA and Size by Decile (1966-2022)

	1	2	3	4	5	6	7	8	9	10
1	16.37***	-	-	-	-	-	-	-	-	-
2	8.28***	10.15***	-	-	-	-	-	-	-	-
3	8.98***	0.98	8.61***	-	-	-	-	-	-	-
4	10.28***	2.64***	1.64	6.35***	-	-	-	-	-	-
5	12.15***	4.98***	3.94***	2.26***	3.41***	-	-	-	-	-
6	12.96***	5.55***	4.34***	2.43***	0.16	5.08***	-	-	-	-
7	14.13***	7.25***	6.04***	4.18***	1.7	2.24***	1.69	-	-	-
8	14.44***	7.80***	6.65***	4.89***	2.54***	3.18***	1.07	0.07	-	-
9	14.84***	8.83***	7.86***	6.37***	4.43***	5.09***	3.39***	2.44***	-2.93***	-
10	21.30***	17.02***	15.93***	14.41***	12.58***	14.37***	12.41***	11.00***	7.16***	-14.47***

The results shown in Figures 4 and 5 and Table 3 provide insight into the overall trend shown in Figure 3. Specifically, they provide evidence consistent with the surmising of Amato and Wilder (1985), based on the theory of Alchian (1965) and Williamson (1963), that there may be a threshold size level, above which growth in size leads to decreasing return on assets. This threshold size level in our study is relative to the year and occurs at approximately the 71st percentile of total assets in a given year. Size increases above this level lead to negative correlation while size decreases below this level lead to positive correlation between ROA and size.

Tests of H3

Moving on to testing how the correlation between ROA and size is related to industry, we first plot the number of SIC 2-digit industries per year and the number of firms per industry-year in Figures 6 and 7 respectively. Notice that most of the SIC 2-digit industries are represented in our sample.[10] Our sample selection criteria of only including industry-year observations with at least five firms is likely the reason why the plot in Figure 6 rises quickly, until the mid-1970s, and then levels off. Compustat's coverage in those earlier years is sparser than in the later years. Notice also how the mean number of firms per industry-year rises up until the dot com bubble and then falls. The median number of firms per industry-year remains stable over time. This points to the fact that there are a few outlier industries with many firms, in the middle years of our sample.

[10] There are 83 SIC 2-digit industries.

Figure 6
# of SIC 2-Digit Industries Per Year (1966-2022)

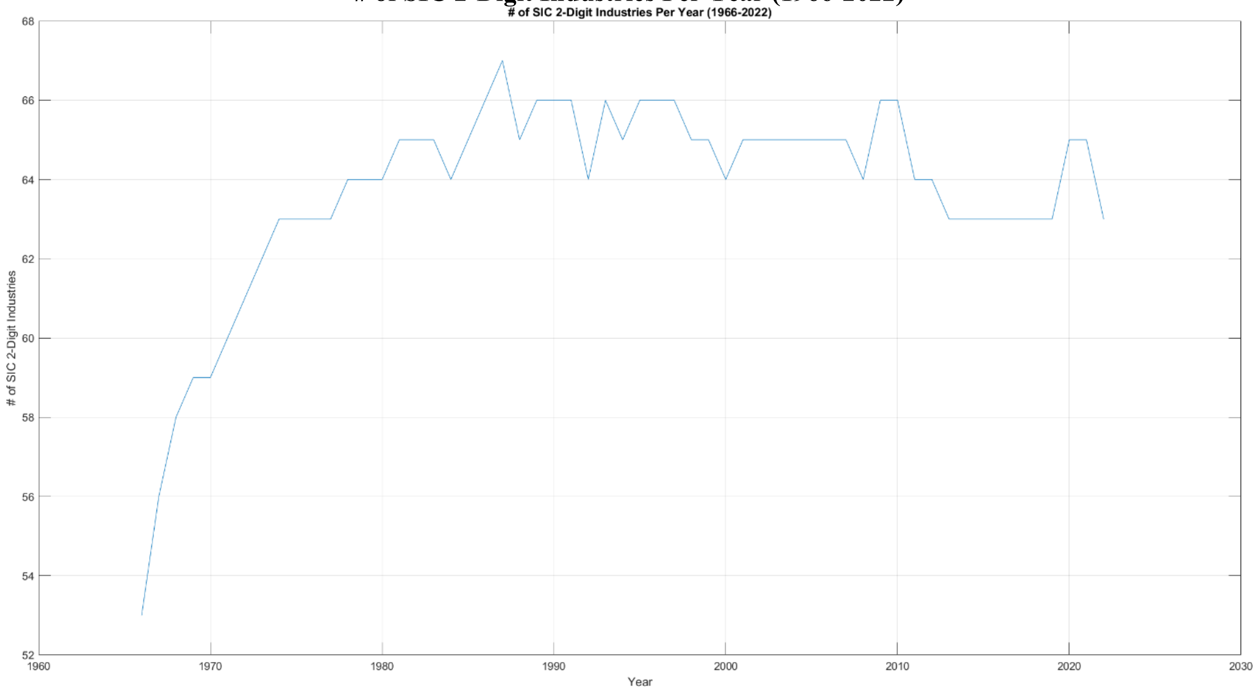
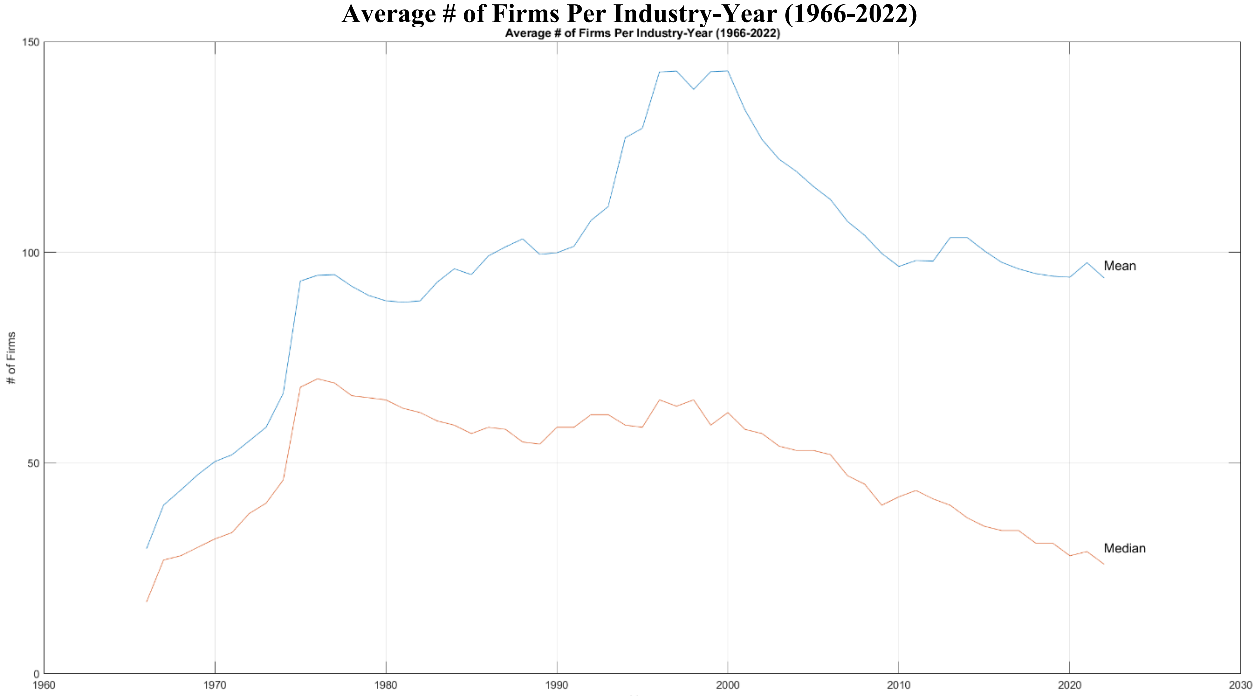


Figure 7
Average # of Firms Per Industry-Year (1966-2022)



In Figure 8 we replicate Figure 3 but require that a given pair of firms be in the same industry. That is, we compute the correlation between the time series of same-industry pair-wise differences in ROA and the time series of the corresponding pair-wise differences in size, for each industry. We then compute the mean and standard deviation of the industry correlations each year and plot these in Figure 8. There are two main takeaways from Figure 8. First, the mean plot looks very similar to Figure 3. So, regardless of whether we compute our correlations within industry or overall, the same trend is observed. Namely, a negative correlation exists between ROA and size from 1966-1974 after which the correlation turns positive and has been increasing from 1975-2022. The Figure 8 plot is shifted downward, relative to the Figure 3 plot as the correlation peaked around 0.45 in the former relative to 0.55 in the latter. Also, the overall positive trend upward is not as stark in the within industry correlation plot. Second, there is considerable variability in the correlation between ROA and size, across industries, in a given year.

Figure 8
Correlation of Pair-wise Difference in ROA and Pair-wise Difference in Log Assets Within-Industry by Year (1966-2022)

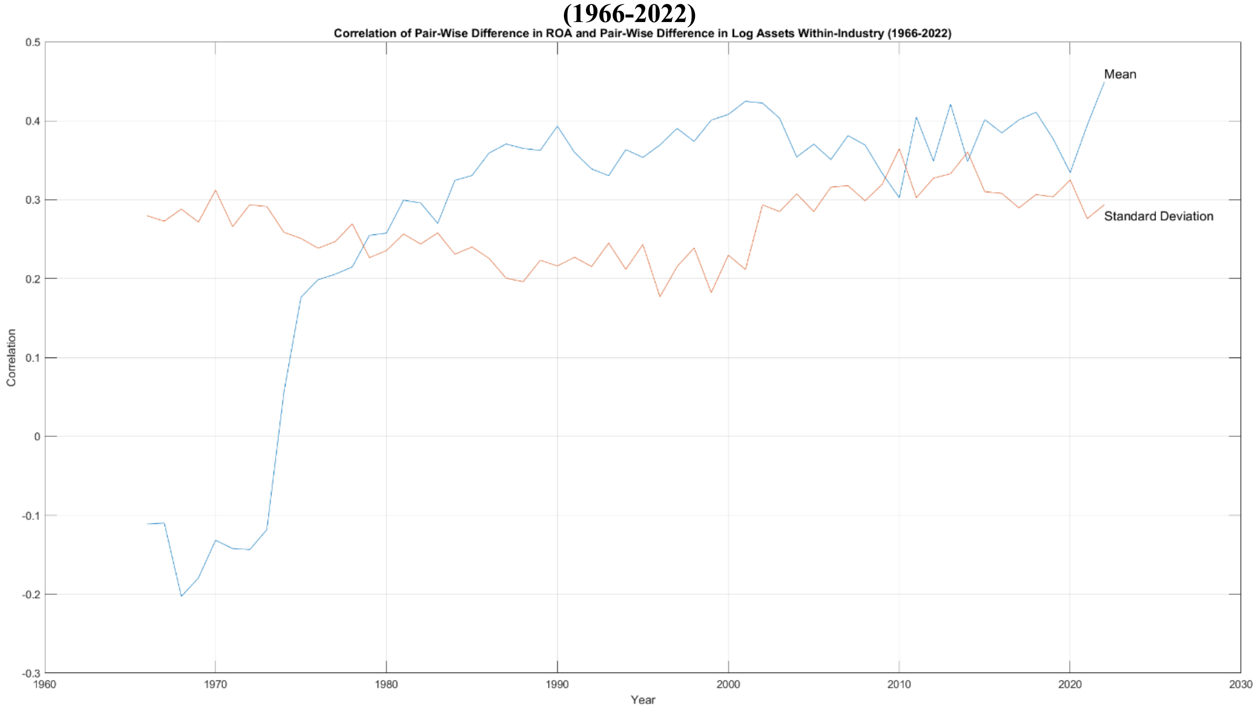
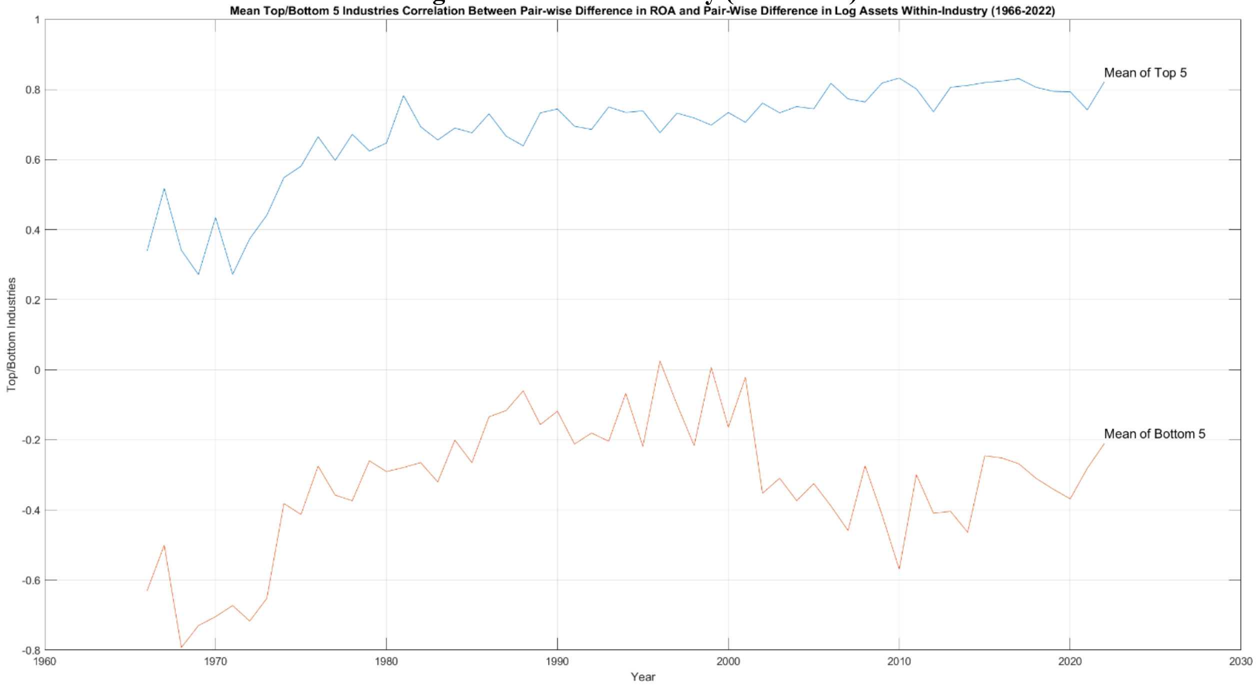


Figure 9 plots the mean correlation values of the five top(most) and bottom(least) correlated industries, over our sample time period.

Figure 9
Mean Top/Bottom 5 Industries Correlation Between Pair-wise Difference in ROA and Pair-wise Difference in Log Assets Within-Industry (1966-2022)



Notice the extremely positive(negative) correlation between ROA and size in the top-5 and bottom-5 industries. Comparing the ROA values between two firms in these industries would not be an apples-to-apples comparison as ROA is strongly correlated with the size of the firm (either positively or negatively) in these industries. The practitioner or researcher should adjust the ROA value in some way or find a better way to compare the profitability of firms in these industries.

Figures 10 and 11 indicate which industries are most often in the top-5(bottom-5) in terms of correlation between ROA and size. For example, SIC 14 -- Nonmetallic Minerals, Except Fuels was in the top-5 in $28/57 \approx 49\%$ of the years while SIC 21 -- Tobacco Products was in the bottom-5 in $37/57 \approx 65\%$ of the years.

Figure 10
SIC 2-Digit Industries in Top-5 of Correlation Between ROA and Size in at Least 20% of the Sample Years

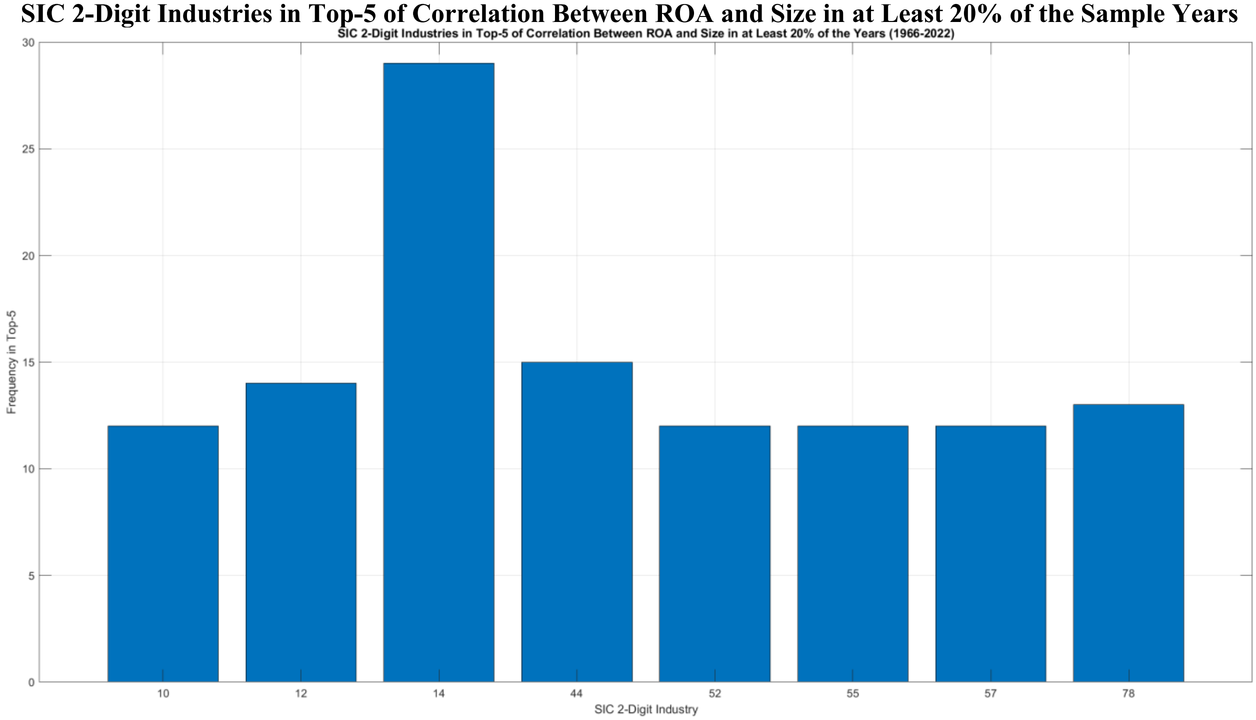
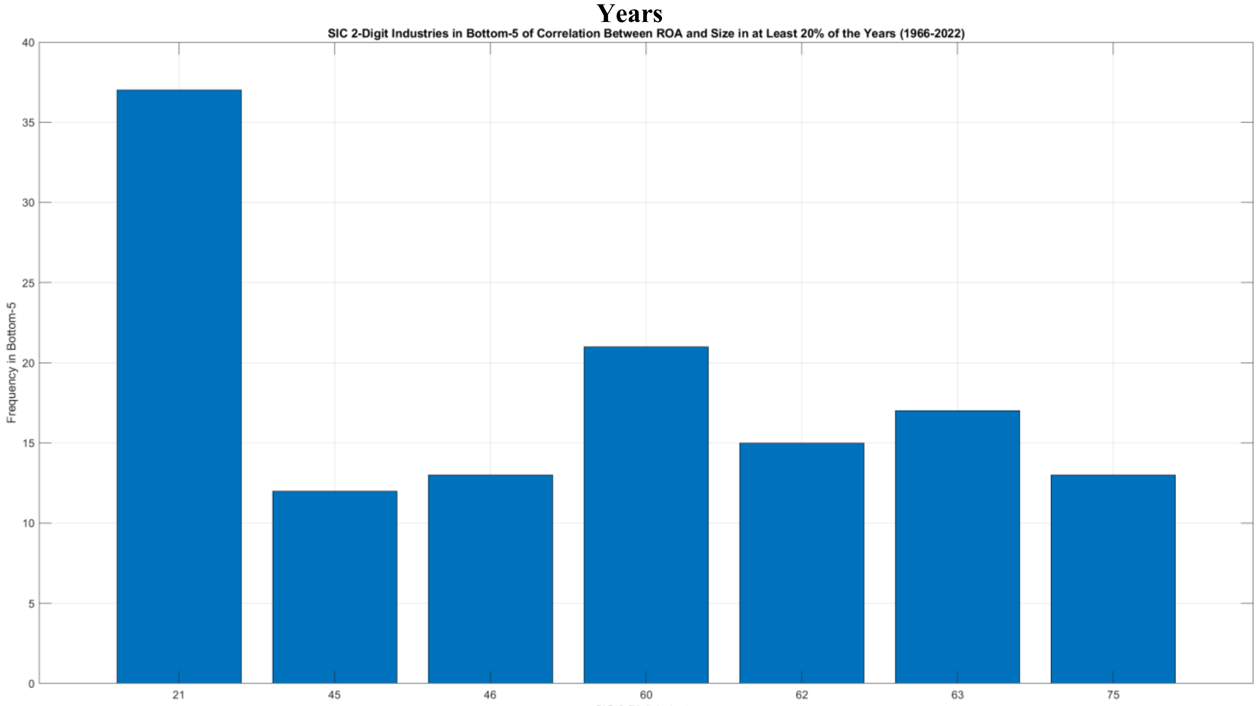
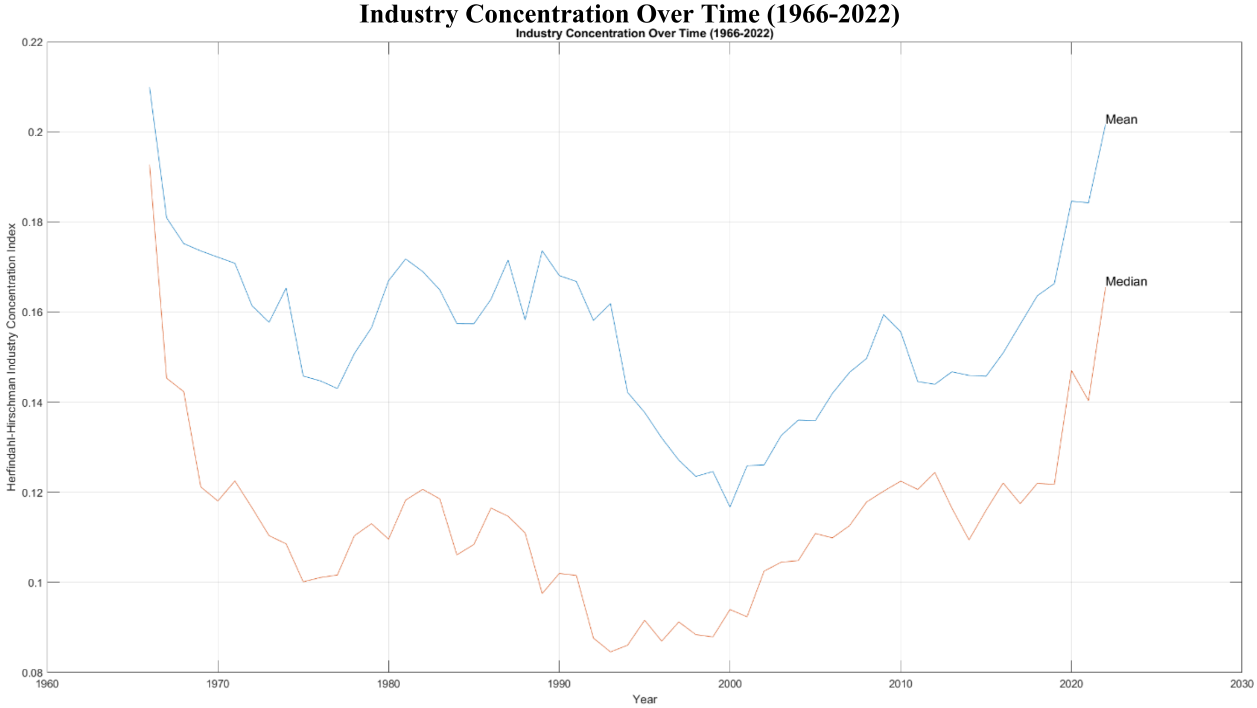


Figure 11
SIC 2-Digit Industries in Bottom-5 of Correlation Between ROA and Size in at least 20% of the Sample Years



Based on our reasoning earlier, regarding the empirical findings of Grullon et al. (2019), we plot the industry concentration over our sample time period in Figure 12. Notice how the industry concentration follows the inverse pattern to the number of firms in the market plot in Figure 2. When there were fewer firms in the market in the late 1960s, industry concentration was at its highest. As the number of firms in the market grew to its all-time high of ≈ 9400 in 1997, industry concentration fell, correspondingly, to its all-time low. The 1994-2014 sub-period of the number of firms and industry concentration plots are consistent with the findings of Grullon et al. (2019) that there are fewer total market firms, fewer firms per industry and industries have become more concentrated and dominated by fewer, bigger firms.

Figure 12
Industry Concentration Over Time (1966-2022)



We directly test H3 by estimating the following linear regression equation,

$$Correlation_{i,t} = \beta_0 + \beta_1 IndConc_{i,t} + \beta_2 IndSize_{i,t} + Year\ Fixed\ Effects + \varepsilon_{i,t} \quad (3)$$

where $Correlation_{i,t}$ is the Pearson correlation between the time series of pair-wise differences in ROA and the time series of corresponding pair-wise differences in log assets, within industry *i* and for year *t*. $IndConc_{i,t}$ is the concentration of industry *i* measured for year *t* using the Herfindahl-Hirschman index.[11] $IndSize_{i,t}$ is the natural logarithm of the number of firms in industry *i* and year *t*.

[11] See Hirschman (1964) for a summary of the index.

We also compute the mean of the dependent and independent variables (across all the industries), for each year, in regression equation (3) and estimate the following regression equation,

$$Correlation_t = \beta_0 + \beta_1 IndConc_t + \beta_2 IndSize_t + \varepsilon_t \quad (4)$$

Table 4 Panel A reports the correlation matrix of the variables in equation (3) and Panel B reports the correlation matrix of the variables in equation (4). Notice how the correlation between the number of firms in the industry and the industry concentration is strongly negative (Panel B) while the correlation between the number of firms in the industry and the ROA with size correlation is strongly positive (Panel B). This implies that leaving out the number of firms in the industry from regression specification (3) or (4) would lead to an omitted variable bias coefficient estimation problem.

Table 4
Correlation Matrices

		$Correlation_{i,t}$	$IndConc_{i,t}$	$IndSize_{i,t}$
Panel A	$Correlation_{i,t}$	1	-	-
	$IndConc_{i,t}$	-0.05	1	-
	$IndSize_{i,t}$	0.068	0.117	1
		$Correlation_t$	$IndConc_t$	$IndSize_t$
Panel B	$Correlation_t$	1	-	-
	$IndConc_t$	-0.407	1	-
	$IndSize_t$	0.788	-0.739	1

Table 5 Panel A reports the results from estimating regression equation (3) and Panel B reports the results from estimating regression equation (4). Notice how the regression results are the same regardless of which specification. Specifically, the coefficient on industry concentration is statistically positive in Panels A (at the 5% level) and B (at the 1% level), holding the number of firms in the industry constant. Notice the relatively high adjusted R^2 values for each regression as well. Table 5 thus provides evidence that is consistent with H3 and thus contributes to the findings of Grullon et al. (2019).

Table 5
Industry Concentration and the Correlation Between ROA and Firm Size

	Variable	Coefficient Estimate
	$Correlation_{i,t} = \beta_0 + \beta_1 IndConc_{i,t} + \beta_2 IndSize_{i,t}$ + Year Fixed Effects + $\varepsilon_{i,t}$	
Panel A	Intercept	-0.200 (-4.59)***
	$IndConc_{i,t}$	0.093 (2.01)**
	$IndSize_{i,t}$	0.024 (4.51)***
	adj(R^2)	30.88%
	N	3,632 industry-years
	Year Fixed Effects	Yes
	$Correlation_t = \beta_0 + \beta_1 IndConc_t + \beta_2 IndSize_t + \varepsilon_t$	
Panel B	Intercept	-6.290 (-10.14)***
	$IndConc_t$	4.175 (3.82)***
	$IndSize_t$	0.685 (12.14)***
	adj(R^2)	77.63%
	N	57 years

CONCLUSION AND FUTURE RESEARCH

The most important overall takeaway from our study is that the traditional accounting metric by which firms have been ranked on performance, ROA, is not without problems. Just because firm A currently earns $5 per dollar of assets while firm B earns $4 per dollar of assets DOES NOT necessitate that firm A is $1 more profitable than firm B, holding size constant. Our study documents that the difference in the two firms ROAs is correlated with their difference in size. The bigger their difference in size the bigger their difference in ROAs (since 1975). In fact, the correlation between the ratio and its denominator has been hovering in the range of [0.5,0.6] over the past 25 years (see Figure 3) and this is alarming. The whole point of dividing by total assets before comparing two firms is to put the firms on the same basis for relative comparison. We document, empirically, that even this fails to do that. ROA itself needs another adjustment before a comparison can be made between the two firms.

We provide evidence that the correlation between ROA and firm size was strongly negative over 1966-1974 and then switched to positive and has been increasingly positive through 2022. Our study is the first, to our knowledge, to document a negative correlation between ROA and size in the distant past and a positive correlation between ROA and size in the past ~50 years in the cross-section of U.S. firms. We find evidence of the approximate inflection point (size threshold), above which further increases in size lead to decreasing income. We find that the correlation between ROA and firm size is negative for the largest firms and positive for the small/medium firms. Furthermore, we document that an approximately zero correlation exists between ROA and firm size for firms that are around the 71st size percentile (based on total assets) in a given year. Our study is the first, to our knowledge, to document these interesting differences in the correlation trend between the extreme small and large decile firms. We shed light on the notion that the smallest firms experience increasing returns to scale while the largest firms experience decreasing returns to scale. We find evidence of the approximate inflection point (size threshold), above which further increases in size lead to decreasing operating income. These findings should be useful to investors, analysts and managers as they seek to predict firm ROA and make decisions accordingly. For example, a manager knowing that their firm is approaching the 71st size percentile (from below) may want to consider downsizing in order that their firm not begin to experience decreasing returns to scale. Also, an analyst or investor can use their knowledge regarding in which size percentile a particular firm falls, to predict future ROA. If the total assets of the firm in question exceed the 71st size percentile then it is likely that ROA will decline in the future, based on our findings. Collectively, our findings contribute to the theory and empirical results first set forth in Amato and Wilder (1985), Alchian (1965) and Williamson (1963), and later, in several foreign empirical studies published in the past 15 years that we discussed on our review of the literature.

We also provide evidence that the correlation between ROA and firm size, within industry, is increasing with the concentration of the industry, holding the number of firms in the industry constant. We thus contribute to the findings of Grullon et al. (2019).

To our knowledge, our study has two limitations. First, we cannot explain what prompted the structural shift in correlation between ROA and size in 1975. Prior to 1975, the correlation, averaged -0.162 and subsequently rose sharply and has averaged 0.401 since that time. Second, we don't provide a way in which ROA can be adjusted so that it no longer is correlated with size, empirically. Future research should examine a way in which to adjust ROA such that the new adjusted ROA is no longer correlated with size, but still captures profitability and thus can be used for proper relative comparison among firms.

The evidence we provide should inform practitioners, educators and researchers when they seek for a way in which to compare the profitability of two firms. Unless ROA is adjusted in some way, it will correlate with size for most firms and thus does not permit an apples-to-apples comparison. Ignoring this will lead to false conclusions when comparing two firms based on ROA. Researchers should also be aware of the correlation between ROA and size when they seek for control variables in their regressions. Including ROA as a control variable for profitability, without also including size, may lead to biased coefficient estimates on the variables of interest.

REFERENCES

Abeyrathna, S. P. G. M., & Priyadarshana, A. J. M. (2019). Impact of Firm size on Profitability. *International Journal of Scientific and Research Publications*, *9*(6), 561-564.

Alchian, A. A. (1965). The basis of some recent advances in the theory of management of the firm. *The Journal of Industrial Economics*, 30-41.

Amato, L., & Wilder, R. P. (1985). The effects of firm size on profit rates in US manufacturing. *Southern Economic Journal*, 181-190.

Babalola, Y. A. (2013). The effect of firm size on firms profitability in Nigeria. *Journal of economics and sustainable development*, *4*(5), 90-94.

Baumol, W. J. (1967). Business behavior, value and growth. New York: *Harcourt Brace Jovanich*

Doğan, M. (2013). Does firm size affect the firm profitability? Evidence from Turkey. *Research journal of finance and accounting*, *4*(4), 53-59.

Becker-Blease, J. R., Kaen, F. R., Etebari, A., & Baumann, H. (2010). Employees, firm size and profitability of US manufacturing industries. *Investment Management and Financial Innovations*.

Grullon, G., Larkin, Y., & Michaely, R. (2019). Are US industries becoming more concentrated?. *Review of Finance*, *23*(4), 697-743.

Hall, M., & Weiss, L. (1967). Firm size and profitability. *The review of economics and statistics*, 319-331.

Hirschman, A. O. (1964). The paternity of an Index. *American economic review*, *54*(5).

Isik, O., Unal, E. A., & Unal, Y. (2017). The effect of firm size on profitability: evidence from Turkish manufacturing sector. *Journal of Business Economics and Finance*, *6*(4), 301-308.

Lev, B., & Sunder, S. (1979). Methodological issues in the use of financial ratios. *Journal of accounting and economics*, *1*(3), 187-210.

Niresh, A., & Thirunavukkarasu, V. (2014). Firm size and profitability: A study of listed manufacturing firms in Sri Lanka. *International journal of business and management*, *9*(4).

Rahman, M. J., & Yilun, L. (2021). Firm size, firm age, and firm profitability: evidence from China. *Journal of Accounting, Business and Management (JABM)*, *28*(1), 101-115.

Kumar, K., Rajan, R., & Zingales, L. (1999). What determines firm size? *National Bureau of Economic Research*

Ravenscraft, D. J. (1983). Structure-profit relationship at the line of business and industry level. *The Review of Economics and Statistics*, 22-31.

Seissian, L. A. (2024). Determinants of internet financial reporting: the case of real estate and development companies in the GCC region. *American Journal of Finance and Accounting*, *8*(1), 28-56.

Smyth, D. J., Boyes, W. J., & Peseau, D. E. (1975). Size, growth, profits, and executive compensation in the large corporation: a study of the 500 largest United Kingdom and United States industrial corporations. London: *Holmes and Meier*

Steindl, J. (1945). Small and big business: Economic problems of the size of firms. London: *Oxford University Press*

Stekler, H. O. (1964). The variability of profitability with size of firm, 1947–1958. *Journal of the American Statistical Association*, *59*(308), 1183-1193.

Susetyo, D. P. (2023). The Effect Of Return On Assets And Firm Size On Capital Structure. *Jurnal Ekonomi, Manajemen dan Akuntansi 1*(01), 25-32.

Williamson, O. E. (1963). Managerial discretion and business behavior. *The American Economic Review*, *53*(5), 1032-1057.