

An Empirical Re-examination of the Relation Between Firm Size and Return¹

Jonathan B. Berk

Department of Finance
School of Business Administration
University of Washington
Box 353200
Seattle, Washington 98195-3200

Tel: (206) 616-8592 Fax: (206) 685-9392

Email: berk@u.washington.edu

Working Paper: 93-BJ-001
March 1993
This Revision: October 9, 1996

Comments Welcomed!

¹ Financial support from The Institute for Quantitative Research in Finance (The Q Group), the Financial Research Foundation of Canada and the Bureau of Asset Management, UBC, is gratefully acknowledged. A copy of this paper, in addition to any other working paper by the author, is available on the authors WWW home page which can be accessed from: <http://weber.u.washington.edu/~bschool/alpha.html>

Abstract

An Empirical Re-examination of the Relation Between Firm Size and Return

Using a data set in which a strong relation between market value and return is known to exist, this paper fails to find a significant relation between average returns and four other measures of firm size. Furthermore, while market value does provide significant additional explanatory power over the single beta model, the same result is not true of the other size measures. The conclusion is that there is no evidence that the size effect is due to the existence of a relation between expected return and firm size. Instead, the results are evidence in favour of the hypothesis that the size effect is due to the endogenous identity relating the market value of a firm to its discount rate.

1 Introduction

A central paradigm in financial economics holds that riskier firms have higher expected returns. Consequently, the large returns earned by stocks with low market values has come to be regarded as evidence that small firms have higher expected returns and are, therefore, subject to certain special risks. However, the lack of a convincing theoretical explanation for why small firms should be riskier has led many researchers to regard the empirical relation between firm size and returns as anomalous.¹ The object of this paper is to examine this interpretation of the size anomaly, that is, that the empirically identified relation between market value and return is evidence that small firms have higher expected returns.

Since its discovery, the observed inverse relation between market value and return has generally been interpreted as evidence that small firms have higher expected returns.² One hypothesis that has been proposed to explain this is that firm size proxies for exposure to a specific risk factor. For example, Chen (1988, p. 183) argues that because “small firms tend to be marginal firms, they fluctuate more with business cycles, and thus, have higher risk exposure to the changing risk premium.” Since small firms have greater exposure, on average, they are riskier, and so there is an inverse relation between the size of firms and returns in the cross-section. Furthermore, if the empirical asset pricing model being tested does not correctly price this factor, then a measure of firm size will add explanatory power to the asset pricing model. The idea that small firms are riskier was first suggested in response to the empirical evidence (in the context of the price-earnings effect) by Ball (1978) who conjectured that earnings proxied for risk.³ This hypothesis has since become the conventional explanation for the empirical relation between market value and average return, or “size effect.”⁴

¹The size anomaly was originally discovered by Banz (1981). The interested reader is referred to any one of the excellent surveys of this literature (e.g. Dimson (1988), Fama (1991) or the special issue of the *Journal of Financial Economics* (vol. 12 no. 1)).

²Banz (1978) himself interpreted the size effect in this way suggesting, in his dissertation, that the effect resulted from the increased costs of diversification associated with smaller stocks.

³Other researchers who provide similar empirically based explanations for the relation between firm size and return include Banz (1978), Roll (1981), Blume and Stambaugh (1983), Schultz (1983), Schwert (1983), Stoll and Whaley (1983), Chan, Chen and Hsieh (1985), Amihud and Mendelson (1986), Keim (1988), Ball and Kothari (1989), Chan and Chen (1991), and Fama and French (1992).

⁴Consequently, Fama and French (1993) have recently run time-series regressions of returns on mimicking portfolios for size because they suggest that “size is associated with a common risk factor that might explain

Recently, Berk (1995) critiqued the above interpretation of the empirical evidence. He argued that since a firm's market value is endogenously determined in equilibrium as the discounted value of expected future cashflows, it depends on the discount rate. For example, if two firms have the same expected cashflow, the one with the larger discount rate will have the lower market value. Consequently, according to this view expected returns will always be negatively correlated with firm market value, *ceteris paribus*. Berk's argument suggests that the traditional interpretation of the empirical relation between market value and average return might be flawed. Rather than evidence of a "size" effect, the relation might be due solely to the endogenous inverse relation between the market value and discount rate of firms.

The object of this paper is to test whether Berk's argument can completely explain the size effect. We therefore take, as the null hypothesis, that there is no relation between the size of a firm and its expected return. The conventional interpretation of the size effect, that firm size is inversely related to expected return, is taken as the alternative hypothesis. Under the null hypothesis, non-market related measures of firm size will not be correlated to average return; while under the alternative, they will be. Consequently, the null hypothesis can therefore be tested by repeating the empirical tests that identified the size effect in the first place with measures of firm size other than market value.

We select four non-market related measures of firm size that are highly correlated to market value: the book value of assets; the book value of un-depreciated, property, plant and equipment; the total value of annual sales; and the total number of employees. Using tests similar to those that uncovered the size effect, we fail to find a significant relation between average returns and any one of these measures of size.

Under the alternative hypothesis the relation between market value and average return is due, in part, to a relation between firm size and expected return. Consequently, if the size of the firm is controlled for, then the predictive power of market value should diminish under the alternative but not under the null. Thus an additional test of the null hypothesis can be constructed by first controlling for the size of the firm and then determining whether

the negative relation between size and average return" (p. 8). They interpret their regression slopes as the equilibrium risk premium of the common risk factor associated with size.

the relation between market value and average return is diminished. We fail to find a diminished relation between market value and return when the size of the firm is controlled for. In light of this evidence, we conclude that it is not possible to reject the null hypothesis that expected return is unrelated to firm size.

Finally, we examine the relation between the firm size and the *risk adjusted* return (that is, the part of return not explained by the CAPM). First, the empirical procedure derived by Gibbons, Ross and Shanken (1989) is used to identify a data set in which market value is strongly related to the risk adjusted return. The same empirical procedure is then repeated with the non-market related measures. None of these measures turn out to be related to the risk adjusted return. The conclusion is that there is no evidence that firm size is responsible for the additional explanatory power of market value over the CAPM.

The rest of the paper is organized as follows. The next section reviews the empirical model that has been used in most size studies. Section 4 then examines the relation between the firm size and return while Section 5 studies the relation between firm size and risk adjusted return, i.e., that part of the return not explained by the single beta model. Section 6 is the conclusion.

2 Empirical Specification

In the standard test of the relation between firm size and return, researchers test whether γ_s is equal to zero in the following relation:

$$E[R_i] = f_i + \gamma_s \text{SIZE}_i \tag{1}$$

where,

$E[R_i]$ = expected return of stock i ,

f_i = expected return specified by an asset pricing model (if it is included in the test) of stock i ,

SIZE_i = a measure of the size of firm i ,

γ_s = coefficient relating firm size to expected return.

Using a variety of empirical procedures, and the logarithm of market value as the measure of firm size, researchers have convincingly demonstrated that γ_s is significantly different from zero.

So long as $f_i \neq E[R_i]$ (i.e., the asset pricing model does not capture all relevant risk factors), Berk (1995) shows that if market value is used as the measure of firm size, then even if expected return and firm size are unrelated, $\gamma_s < 0$.⁵ He therefore argues that the empirical results relating market value and return cannot be used as evidence that the size of a firm is related to its expected return. To show that size and return are related, it is necessary to show that γ_s is different from zero even when one employs measures of firm size that are not market related.

In this paper we first repeat, using four non-market size measures, the form of the test that does not include a specific asset pricing model: Equation (1) is tested with $f_i \equiv a$, where a is any constant. The relation between firm size and risk adjusted return is then studied. This entails repeating existing studies in which (1) is tested with $f_i = \hat{E}_i$, where \hat{E}_i is the expected return predicted by the single beta model.

3 The Data

The data set consists of all NYSE firms for which data exists in the period July 1966 - June 1987. This data set is chosen for two reasons. First, it closely resembles the data set in which Fama and French (1992) detected a very large size effect. The object is to determine whether the size effect results, in part, from a relation between firm size and return. Thus it is essential that the data set used is one in which the size effect has already been empirically demonstrated to exist. Secondly, restricting attention to NYSE stocks avoids clouding the inferences of the tests by including new firms. New firms are on average small, and one could argue, riskier.⁶ If this is true, then they would have higher expected returns. But

⁵The actual assumptions required are that expected return, $E[R_i]$, and the prediction of the asset pricing model, f_i , be uncorrelated to expected cashflow.

⁶The idea that new firms are necessarily riskier is by no means obvious. For instance Clarkson, Guedes and Thompson (1993) provide conditions when new firms will not be riskier.

in this case, it is the firm's short life rather than its size that is associated with its higher expected return. Since a very strong size effect has been identified on the NYSE alone, and since few firms listed on the NYSE are new firms, limiting attention to firms on this stock exchange provides a more powerful test of the null hypothesis.

All return and market value data are obtained from the monthly CRSP return and price files. Four non-market measures of the size of the firm are used: (1) the book value of assets (BVA); (2) the book value of all, un-depreciated, property, plant and equipment (PPE); (3) the total value of annual sales; and (4) the total number of employees. These data are obtained from the COMPUSTAT annual industrial files of income statement and balance-sheet data.⁷

The above four measures of the size of the firm are natural choices because there is no obvious reason why any of them should contain an adjustment for risk. The two accounting measures (BVA and PPE) might neglect an important component of firm size, the value of intangible assets. The intangible assets of a firm are defined as the difference between the market capitalization and the book value of assets. Therefore, the risk premium is part of the intangible assets of the firm. Yet there are other intangibles besides the risk premium (e.g. patents, copyrights, etc.) that can be thought of as part of the size of the firm but would not be measured by accounting variables. It is hard to see why these assets would have a significantly different relation to the return of the firm than the tangible assets, but if they do, the accounting measures would misstate the relation between the size of the firm and return. For this reason non-accounting measures of firm size are included (i.e., annual sales and number of employees).

Despite their differences, all five measures are, cross-sectionally, highly correlated. (See Table 1 which reports the time-series averages of these correlations.) The correlations of the other four size variables with market value are lower than their correlations with each other.

A selection bias in this data set merits comment. The data set was intentionally selected because a strong relation existed between market value and average return. As such, the analysis is subject to the kind of data snooping bias discussed by Denton (1985), Black

⁷The COMPUSTAT identification numbers are 6,7,12 and 29 respectively.

Table 1: Average Cross-Sectional Correlation Coefficients and Standard Deviations of the Five Size Measures: 1966-1986

Each average correlation coefficient is the average of the monthly cross-sectional correlation coefficients of the five size variables. The average standard deviation (the last row in the table) is the average of the monthly cross-sectional standard deviations of the five size variables. The market value of equity (MV) is measured in June of year t . Book value of assets (BVA), value of property, plant and equipment (PPE), annual sales (sales) and number of employees (employees) are the reported numbers for firms with fiscal year ends from June of year $t - 1$ to May of year t . With the exception of the number of employees, all units are billions of dollars. Number of employees is quoted in thousands.

	MV	BVA	PPE	Sales	Employees
MV	1				
BVA	.732	1			
PPE	.721	.938	1		
Sales	.692	.822	.767	1	
Employees	.683	.759	.703	.778	1
Standard Deviation	2.377	3.988	3.563	3.782	41.39

(1992), Foster and Smith (1992), Lo and MacKinlay (1992) and others. As these authors point out, if classical statistical significance tests are used, this bias increases the probability of rejecting the null in the sample when it holds in the population.

The actual matching of the accounting and market data is accomplished as follows. The COMPUSTAT file in year t contains data for fiscal year ends from June of year t to May of year $t + 1$. In order to guarantee that the relevant explanatory variables are known before the returns they are supposed to explain, accounting data in the year t COMPUSTAT file is matched with the market value at the end of June in year $t + 1$. These variables are then used to explain returns from July of year $t + 1$ through June of year $t + 2$.⁸ As a result one year of data is lost leaving a total of 20 years (or 240 months) of observations.

The one-month lag between the last possible fiscal year end (May) and the first return observation (July) is necessary because firms report their balance sheet data after their fiscal year ends. Some might argue that this lag is not long enough, since a small minority of firms

⁸Thus, the accounting numbers could be at most a year behind the market value number. This disparity is not, however, a cause for concern. Using market values from June of year $t - 1$ (i.e., ensuring that market value is never more current than any accounting number) only marginally changes the results and does not change any inference in the paper.

do not report within even the legally mandated 90 day period. However, less than 10% of firms have fiscal year ends between February and May and so only a tiny minority of firms' accounting data would be unavailable as late as July. The cost in potential explanatory power was therefore not deemed worth the benefits obtained by increasing this lag.

4 The Relation Between Size and Return

The purpose of this section is to determine whether any part of the inverse relation between the market value and return of stocks listed on the NYSE can be attributed to a relation between firm size and expected return. To get a sense of the effect, we follow Fama and French's (1992) lead and first construct a set of informal, expository tests based on the average returns of portfolios sorted by the different size measures. The null hypothesis is then formally tested using the same test Fama and French (1992) used.

4.1 Informal Tests⁹

In June of each year, using book value of assets and market value individually, we sort the stocks into quintiles of decreasing size. We then measure the average return of each quintile over the following year (i.e., July through June). Table 2 (in the *columns* labelled "All") contains the average of these averages for each size variable, that is, the average return of each quintile over the sample period. The returns of the quintiles sorted on market value (Panel B) monotonically increase as the size of the quintile decreases. Not surprisingly, these results mirror the results obtained by Fama and French (1992, Table I). The results for the quintiles sorted using BVA are reported in Panel A in each table. The largest size quintile always has a lower average return than the smallest size quintile and, for the most part, the larger the quintile the lower the return. However, the largest quintile does not contain the smallest return and the dispersion across the quintiles is clearly not as large as it is for market value.

Under the null hypothesis, the correlation between average return and market value should not be reduced when the size of the firm is controlled for. Therefore, if stocks are

⁹We report the results for informal tests using the book value of assets only. The results for the other three non-market size variables are not qualitatively different and are available from the author on request.

Table 2: Average Monthly Returns (in %) of Portfolios Formed on Market Value of Equity (MV) and the Book Value of Assets (BVA): 1966-1987

Each June firms are sorted into 5 quintiles using BVA. Each quintile is then sorted into 5 portfolios based on the MV. The average return of each portfolio over the next 12 months is then recorded. Each cell in the following table is the average of these averages over the period July 1967 - June 1987. The column (row) labeled “All” is the average across the cells in each respective row (column). The market value of equity (MV) is measured in June of year t . Book value of assets (BVA) is the reported number for firms with fiscal year ends from June of year $t - 1$ to May of year t .

Panel A: Sort Order — BVA and then MV						
BVA Quintiles	MV Quintiles					All
	largest	4	3	2	smallest	
largest	.95	.98	1.07	1.23	1.47	1.14
4	1.08	1.02	1.10	1.09	1.37	1.13
3	.95	1.15	1.20	1.47	1.55	1.26
2	.95	1.25	1.29	1.42	1.78	1.34
smallest	1.09	1.27	1.43	1.50	2.05	1.47
All	1.00	1.13	1.22	1.34	1.65	

The table below is constructed similarly except the order of the sorts is reversed: 5 MV based quintiles are formed and each quintile is then sorted into 5 BVA based portfolios.

Panel B: Sort Order — MV and then BVA						
MV Quintiles	BVA Quintiles					All
	largest	4	3	2	smallest	
largest	1.01	.94	.95	1.13	.93	.99
4	1.30	1.20	.98	.91	1.02	1.08
3	1.35	1.10	1.36	1.20	1.22	1.25
2	1.40	1.48	1.35	1.23	1.22	1.34
smallest	1.81	1.64	1.61	1.69	1.61	1.67
All	1.38	1.27	1.25	1.23	1.20	

first sorted into quintiles based on a size variable other than market value, under the null hypothesis, sorts on market value within each quintile should produce the same results as the sort by market value on the entire data set. Tables 2 (Panel A) provides the average return of 25 portfolios formed by first sorting into quintiles by BVA, and then sorting, within each quintile, by market value. The dispersion in average returns within any BVA quintile is shown in the rows of Panel A. The bottom row of this panel provides the average return of each market value quintile (i.e., the variable sorted last). It therefore provides a measure of the overall effect of market value when controlling for the effect of BVA. In every case market value continues to explain average return. Furthermore, the dispersion in average returns appears to be no different than when stocks are sorted on market value alone, and this dispersion is still much greater than BVA.

If the order of the sorts is reversed the results differ dramatically (see Table 2, Panel B). Once the market value of the firm is controlled for, the size of the firm appears to be *positively* related to average return. Within each MV quintile, larger BVA seems to be associated with larger, not smaller, average returns. Although at first glance surprising, on further reflection this result is consistent with the null hypothesis. To see why, recall that the market value of a firm is influenced by two effects: larger firms have larger market values and firms with high discount rates have lower market values. Consider then the set of firms in the lowest market value quintile. These firms consist of both small firms and slightly larger firms that are in this quintile because they have higher discount rates. That is, the only way a larger firm could get into this quintile is if it has a large discount rate. Consequently, *within this quintile*, the larger the firm the higher the discount rate. Similarly, the only way the smallest firm in the largest market value quintile could remain in that quintile is if it has a small discount rate. Thus the smaller the firm within this quintile the lower the discount rate. When firm size itself is unrelated to the discount rate, within any market value quintile firm size should be associated with larger average returns.

Although the sorts appear to reveal that even when controlling for the size of the firm, MV does a better job explaining the cross-sectional distribution of average returns than the other size measures, it is unclear whether this difference is statistically significant. This question is addressed in the following section.

4.2 Formal Tests

Although the results of the informal tests are illuminating, no inferences on whether the null can be rejected can be made without a formal test. Therefore, in this section a standard Fama-Macbeth regression test of the null hypothesis that $\gamma_s = 0$ in (1) (when f_i is taken to be constant across assets) is undertaken.

Table 3 contains the time series average slopes (in percent) of the month by month Fama-Macbeth (FM) univariate cross-sectional regressions of the monthly return of all firms on the logarithm of each of the five size measures separately. The associated t -statistic is the

Table 3: Average Slopes (in %) and t -statistics from Monthly Regressions of Stock Returns on Five Size Measures: July 1967 - June 1987

The regressions are month by month cross-sectional Fama-Macbeth (FM) univariate regressions of the monthly return of all firms on the logarithm of each of the five size measures separately. The coefficient is the average of the monthly regression slopes and the t -statistic is this average divided by its time series standard error. The p -values are two tailed and are calculated under the assumption that the t -statistic is normally distributed. The market value of equity (MV) is measured in June of year t . Book value of assets (BVA), value of property, plant and equipment (PPE), annual sales (sales) and number of employees (employees) are the reported numbers for firms with fiscal year ends from June of year $t - 1$ to May of year t . The returns are measured monthly from July of year t to June of year $t + 1$. Thus each monthly return, from July of year t to June of year $t + 1$ is regressed onto size measures that are recorded on or before June of year t .

	MV	BVA	PPE	Sales	Employees
Coefficient (%)	-0.170	-0.099	-0.082	-0.069	-0.061
t -statistic	-2.50	-1.83	-1.52	-1.47	-1.68
p -value	.0124	.0671	.129	.142	.0922

standard FM t -statistic (the average slope divided by its time series standard error). The p -value represents the probability, under a normal distribution, of observing a t -value greater (in absolute magnitude) than the value actually observed. Since the dependent variables are all directly observable, there is no error-in-variables problem, and so, given the length of the sample period, the t -statistics should be normally distributed under the null hypothesis. This was confirmed by bootstrapping the distribution.

Table 3 verifies that market value is a significant predictor of future return. The results for the other four size measures differ, however. Because of the data snooping bias dis-

cussed in Section 3, classical p -values overstate the significance of the other size variables.¹⁰ Normally the test would need to be adjusted to take this bias into account. However, since the null hypothesis cannot be rejected at even the classical cutoffs, adjusting for the data snooping bias is unnecessary, since it cannot change the inference.¹¹

Table 4: Average Cross-Sectional Correlation Coefficients of the Residuals of the Logarithm of Market Value Regressed onto the Logarithm of the Four Size Measures: 1967-1986

Each average correlation coefficient is the average of the monthly cross-sectional correlation coefficients of the residuals. The coefficients are listed under the size variable from which the residual is calculated. The market value of equity (MV) is measured in June of year t . Book value of assets (BVA), value of property, plant and equipment (PPE), annual sales (sales) and number of employees (employees) are the reported numbers for firms with fiscal year ends from June of year $t - 1$ to May of year t .

	BVA	PPE	Sales	Employees
PPE	.832			
Sales	.775	.710		
Employees	.647	.638	.883	
All	.960	.847	.881	.762

The null hypothesis that $\gamma_s = 0$ cannot be rejected based on the evidence from the above FM regression test. However, one might be concerned that the above test might not be powerful enough to reject the null. One way to address this concern is to compare the relation between market value and return when firm size is controlled for and when it is not. Under the alternative, if firm size accounts for any part of return, the predictive power of market value should decrease when firm size is controlled for.

To control for firm size, the logarithm of market value is annually cross-sectionally regressed onto the logarithm of each size measure individually. Log market value is also annually cross-sectionally regressed onto the log of all four size variables together. The

¹⁰The bias follows because a strong relation between market value and return was known to exist in the data sample prior to the study. Since the non-market related measures are known to be highly correlated with market value, under the null hypothesis, this bias increased the probability of detecting an effect in the sample when none exists in the population (see Lo and MacKinlay (1990) for a discussion of this bias in the context of the size effect).

¹¹There is no evidence of heteroscedasticity in any of the cross-sectional regressions. As such, the results are unaffected when the residual covariance matrix is estimated using White's heteroscedastic-consistent estimator.

residuals of these regressions control for firm size because they are the part of market value that is orthogonal to each size measure respectively. They are consequently highly correlated (see Table 4).

Table 5: Average Slopes (in %) and t -statistics from Monthly Univariate Regressions of Stock Returns onto Orthogonalized Market Value Measures: July 1967 - June 1987

Each year the logarithm of every firm's market value is cross-sectionally regressed on the logarithm of each of the four size measures individually producing four residuals. A fifth residual is created by cross-sectionally regressing the logarithm of every firm's market value onto the logarithm of *all* four size measures together. Over the next year, the return of every firm in a given month is then cross-sectionally regressed onto these five residuals individually. The coefficient of the residual regressions is the average of these monthly regression slopes and the t -statistic is this average divided by its time series standard error. The results are reported under the size measure from which the residuals are calculated. The p -values are two tailed, and are calculated under the assumption that the t -statistic is normally distributed. The market value of equity (MV) is measured in June of year t . Book value of assets (BVA), value of property, plant and equipment (PPE), annual sales (sales) and number of employees (employees) are the reported numbers for firms with fiscal year ends from June of year $t - 1$ to May of year t . The returns are measured monthly from July of year t to June of year $t + 1$. Thus each monthly return, from July of year t to June of year $t + 1$ is regressed onto orthogonalized size measures that are recorded on or before June of year t .

	Part of Market Value Orthogonal to:				
	BVA	PPE	Sales	Employees	All
Coefficient (%)	-0.244	-0.237	-0.295	-0.240	-0.280
t -statistic	-2.65	-2.90	-3.29	-2.67	-2.71
p -value	.0081	.0037	.0010	.0076	.0068

The relation between these residuals and average return is then compared to the relation between market value and average return. Under the alternative this relation should diminish (γ_s should be closer to zero) when these residuals, rather than market value, are used as the measures of firm size. The above FM univariate cross-sectional regression test is therefore repeated using each residual individually as the independent variable. The coefficients of these regressions are reported in Table 5, and can be found under the size variable from which the residual is calculated. The coefficients are all negative and significantly different from zero. More importantly, the coefficient of every residual is *larger* (in absolute value) than the coefficient of the market value variable itself.

Finally, we can also measure the marginal contribution, over market value, of the other

size variables by repeating the above procedure using the residuals of the cross-sectional regression of each size variable on market value. The results of this test are reported in Table 6. They reflect the informal evidence in Table 2. Controlling for market value, γ_s actually reverses sign. In addition, the difference in the coefficient estimates in Tables 3 and 6 is significant.¹²

Table 6: Average Slopes (in %) and t -statistics from Monthly Univariate Regressions of Stock Returns onto Size Measures Orthogonal to Market Value: July 1967 - June 1987

Each year the logarithm of each of the four size measures is individually cross-sectionally regressed on the logarithm of every firm's market value producing four residuals. Over the next year, the return of every firm in a given month is then cross-sectionally regressed onto these four residuals individually. The coefficient of the residual regressions is the average of these monthly regression slopes and the t -statistic is this average divided by its time series standard error. The results are reported under the size measure from which the residuals are calculated. The p -values are two tailed, and are calculated under the assumption that the t -statistic is normally distributed. The market value of equity (MV) is measured in June of year t . Book value of assets (BVA), value of property, plant and equipment (PPE), annual sales (sales) and number of employees (employees) are the reported numbers for firms with fiscal year ends from June of year $t - 1$ to May of year t . The returns are measured monthly from July of year t to June of year $t + 1$. Thus each monthly return, from July of year t to June of year $t + 1$ uses orthogonalized size measures that are recorded on or before June of year t .

	Orthogonalized to Market Value:			
	BVA	PPE	Sales	Employees
Coefficient (%)	0.091	0.079	0.163	0.093
t -statistic	1.39	1.32	3.16	1.76
p -value	.165	.186	.00157	.0782

As was pointed out in Section 4.1, a positive relation between firm size measures and return controlling for risk is consistent with the hypothesis that firm size and risk are unrelated. It is tempting to use the statistical significance of the marginal explanatory power of sales as further evidence in favor of the hypothesis that size and risk are unrelated. A note of caution is in order. One would expect such an effect to be pervasive and so it

¹²A two-tailed t -test is used to test whether the coefficient estimate of the FM regression of return on a particular size measure is different to the coefficient estimate of the FM regression of return on the residuals of the same size measure on market value (that is, the t -statistic is defined as the mean of the monthly differences between the two slopes divided by the standard error of these differences). The t -statistics are 2.59, 2.92, 3.34 and 2.62 with corresponding p -values of 0.0095, 0.0035, 0.0008 and 0.0087 for BVA, PPE, sales and employees respectively.

should be reflected in all four measures, not just sales. Further, in a study where so many significance tests are conducted one should be careful to draw an inference based on any one significant test.

5 The Relation Between Size and Beta

One might argue that the empirical literature on size is not only regarded as anomalous because it finds that market value is an inverse predictor of return, but also because when the single beta model and market value are tested simultaneously, market value has additional explanatory power.¹³ Although Fama and French (1992) interpret this evidence as “a shot strait at the heart (p. 438)” of the CAPM, this result by itself cannot be used to reject the CAPM (see Berk (1995)). So long as there is any misspecification in the empirical test, it is possible to reject the CAPM in favour of market value even if the CAPM itself holds perfectly. The same result is not necessarily true of the other size measures. If the size of the firm is unrelated to its return, then even if the CAPM does not hold perfectly, the size of the firm should have no additional explanatory power in a joint test of it and the CAPM. The purpose of this section is to empirically verify this fact.

Formally, the null hypothesis is that, in (1), $\gamma_s = 0$ when $f_i = r + \beta_i \bar{R}_m$, where β_i is the CAPM beta of stock i , r is the riskless rate of interest and \bar{R}_m is the expected return of the market proxy. The alternative hypothesis is that firm size has additional explanatory power, that is, that $\gamma_s \neq 0$. Using market value as the measure of firm size many researchers have been able to reject this null. Perhaps the most convincing evidence can be found in Gibbons, Ross and Shanken (1989). The reason why this evidence is particularly strong is that the empirical procedure they use does not require measuring β_i and so their test does not contain an error-in-variables bias. In this section we therefore repeat their procedure, in a data set in which a strong relation between market value and risk adjusted return is known to exist, using other measures of firm size.

The methodology that Gibbons, Ross and Shanken (1989) use is based on the time series test of the CAPM originally developed by Black, Jensen and Scholes (1972) and extended

¹³See, e.g., Jegadeesh (1992) or Fama and French (1992)

to include a test of the size effect by Brown, Kleidon and Marsh (1983). Unfortunately, as Gibbons, Ross and Shanken (1989) show in their paper, over the time period in this paper, there is no significant relation between market value and risk adjusted return on the NYSE.¹⁴ Most researchers would consider finding no relation between the non-market size measures and risk adjusted return in a data set in which no relation between market value and risk adjusted return was known to exist as uninteresting. It therefore makes little sense to use this sample in this study. It turns out, however, that if the SP500 stocks are removed from the sample, a strong relation between market value and risk adjusted return does exist. Consequently, all the empirical tests in this section will be restricted to the NYSE stocks that are followed by COMPUSTAT but are not in the SP500 index.

In June of each year, each size measure is individually used to sort the sample of stocks into 10 size-ranked portfolios. The excess return of each decile over the next 12 months is then recorded.¹⁵ A 240 month time series of monthly average excess returns for each decile is then formed by concatenating these returns over the 20 year period July 1967 - June 1987. Each time series is then separately regressed onto the time series of monthly excess returns of the CRSP value weighted index. The null hypothesis that $\gamma_s = 0$ is equivalent to the condition that the intercepts of the decile regressions are all jointly zero.

Table 7 contains the regression intercepts for each decile. They are also graphed in Figure 1. The estimates increase approximately monotonically as the market value of the portfolio decreases. Furthermore, of the ten t -statistics of the portfolios sorted by market value, three are above 1.96. The corresponding results for the other four size measures are not as convincing. Of the 40 t -statistics of all the portfolios sorted by the other four size measures only two are above 1.96. However, as Gibbons, Ross and Shanken (1989) point out, the t -statistics do not provide conclusive evidence on whether all the intercepts together are significantly different from zero.

Table 8 reports the results of the F -test that the intercepts are jointly significantly different from zero. The results show that when either market value or book-to-market is

¹⁴Gibbons, Ross and Shanken (1989) divided the time period in this paper into two subperiods and report the results for each subperiod. The comparable W -statistic for the whole time period is 0.069 which corresponds to a p -value of .11.

¹⁵The excess return of each stock is calculated by subtracting the 1 month riskless rate in the CRSP bond file from the monthly return of each asset.

Table 7: The Mean-Variance Efficiency of the CRSP Value Weighted Index: Univariate Test

Using each size measure as well as book-to-market, the firms are (separately) sorted into 10 sized based deciles in June of each year. That is, each June, stocks are sorted into deciles based on the observed market value of equity (MV) in June of year t and on the reported values between June of year $t - 1$ and May of year t , of the book value of assets (BVA), the value of property, plant and equipment (PPE), annual sales (sales) and the number of employees (emp). The monthly excess return (i.e., the realized return minus the CRSP one month riskless rate) of each decile from July of year t to June of year $t + 1$ is then recorded. A 20 year time series of monthly returns for each decile is then formed by concatenating the 12 monthly returns for each decile in each year over the period July 1967 - June 1987. Each of these time series are then separately regressed onto the time series of monthly excess returns of the CRSP value weighted index. The intercept (in %) of each regression and its associated t -statistic are recorded in the table.

Size Dec	MV		BVA		PPE		Sales		Emp	
	$\hat{\alpha}_{im}$	t	$\hat{\alpha}_{im}$	t	$\hat{\alpha}_{im}$	t	$\hat{\alpha}_{im}$	t	$\hat{\alpha}_{im}$	t
lg	-0.19	-2.12	0.17	1.26	0.13	1.03	0.10	0.78	0.16	1.01
9	-0.11	-1.10	0.06	0.40	0.17	1.32	0.10	0.64	0.10	0.61
8	-0.06	-0.45	0.05	0.34	0.17	1.08	0.20	1.20	0.14	0.74
7	0.02	0.15	0.19	1.16	0.22	1.21	0.21	1.20	0.17	0.88
6	0.21	1.25	0.16	0.90	0.24	1.29	0.19	0.97	0.21	1.10
5	0.12	0.65	0.19	0.91	0.22	1.09	0.13	0.68	0.14	0.73
4	0.42	1.91	0.33	1.64	0.19	0.92	0.23	1.16	0.17	0.88
3	0.27	1.13	0.16	0.69	0.31	1.33	0.27	1.32	0.33	1.83
2	0.56	2.10	0.40	1.73	0.24	1.02	0.42	1.83	0.46	2.58
sm	0.83	2.45	0.43	1.69	0.36	1.33	0.31	1.26	0.40	2.09

Figure 1: Intercepts of the Univariate Test for Efficiency of the CRSP Value Weighted Index.

Using each size measure, the firms are (separately) sorted into 10 sized based deciles in June of each year. The average excess return (i.e., the realized return minus the CRSP one month riskless rate) of each decile over the next 12 months is then recorded. A time series of monthly returns for each decile is then formed by combining these annual returns over the 20 year period July 1967 - June 1987. These time series are then separately regressed onto the time series of monthly excess returns of the CRSP value weighted index. The graph shows the intercept terms of each regression.

used to form the portfolios, the intercepts are significantly different from zero. However, the other four size measures have no detectable explanatory power whatsoever. None of the other effects are significant (BVA had the smallest $p = 0.08$) at even classical statistical cutoffs.

In a data set in which a strong relation between market value and return exists, there is no evidence that the size of the firm has any additional explanatory power over and above the single beta model. In short, even when the proxy portfolio is shown to be inefficient, there is no evidence that firm size has additional explanatory power.

Table 8: The Mean Variance Efficiency of the CRSP Value Weighted Index: Multivariate Test

Using each size measure, the firms are (separately) sorted into 10 sized based deciles in June of each year. That is, each June stocks are sorted into deciles based on the observed market value of equity (MV) in June of year t and on the reported values between June of year $t - 1$ and May of year t , of the book value of assets (BVA), the book value of common equity, the value of property, plant and equipment (PPE), annual sales (sales) and the number of employees (employees). The monthly excess return (i.e., the realized return minus the CRSP one month riskless rate) of each decile from July of year t to June of year $t + 1$ is then recorded. A 20 year time series of monthly returns for each decile is then formed by concatenating the 12 monthly returns for each decile in each year over the period July 1967 - June 1987. Each of these time series are then separately regressed onto the time series of monthly excess returns of the CRSP value weighted index. The intercepts are then used to calculate the Gibbons, Ross and Shanken W statistic. Under the null hypothesis, $\frac{T(T-N-1)}{N(T-2)}W$ has a central F distribution with degrees of freedom N and $T - N - 1$. Thus the p -values represent the probability, under a F distribution with degrees of freedom 10 and 229, of observing a (transformed) W value greater than the observation actually observed.

	MV	BVA	PPE	Sales	Employees
W	0.149	0.074	0.013	0.029	0.066
p -value	.00031	.082	.980	.755	.129

6 Conclusion

We fail to find evidence to support the hypothesis that there is a cross-sectional relation between firm size and expected return. In particular, in a data set in which market value has strong cross-sectional relation to average return, this paper fails to find a similar significant relation between average return and other, non-market, measures of firm size. We show

that the relation between market value and average return is unaffected when non-market measures of firm size are used to control of the size of the firm. On the other hand, when the market value of the firm is controlled for, we find evidence of a *positive* relation between the non-market size variables and average return. We interpret these findings as evidence that the cross-sectional explanatory power of market value does not result from a relation between firm size and return.

The empirical results in this paper explain why Fama and French (1992) find that book-to-market is more significantly related to return than market value is. The large cross-sectional variation in firm size adds statistical noise to market value which weakens the significance of the relation between market value and return. Book-to-market, however, controls for firm size and so is not subject to this problem.

Finally, in a data set in which market value provides significant additional explanatory power over and above the single beta model, this paper demonstrates that other, non-market, measures of firm size fail to provide similar explanatory power. Taken together, the results in this paper are evidence against the hypothesis that firm size proxies for exposure to specific risk factors.

References

- Amihud, Y. and H. Mendelson (1986), "Asset pricing and the bid-ask spread," *Journal of Financial Economics*, **17**:223-250.
- Ball, R., (1978), "Anomalies in relationships between securities' yields and yield-surrogates," *Journal of Financial Economics*, **6**:103-126.
- Ball, R., and S.P. Kothari (1989), "Nonstationary expected returns: Implications for tests of market efficiency and serial correlations in returns," *Journal of Financial Economics*, **6**:103-126.
- Banz, R.F., (1978), "Limited Diversification and Market Equilibrium: An Empirical Analysis," Ph.D. Dissertation, University of Chicago, Chicago, Ill.
- Banz, R.F., (1981), "The relation between return and market value of common stocks" *Journal of Financial Economics*, **9**:3-18.
- Berk, J.B. (1995) "A Critique of Size Related Anomalies," *Review of Financial Studies*, **8**:275-86.
- Black, F., (1992), "Beta and Return", presentation at The Berkeley Program in Finance: Are Betas Irrelevant? Evidence and Implications for Asset Management.
- Black, F., M.C. Jensen and M. Scholes (1972), "The Capital Asset Pricing Model: Some Empirical Tests." in Michael C. Jensen ed.: *Studies in the Theory of Capital Markets*, Praeger Publishers, New York, USA.
- Blume and Stambaugh (1983), "Biases in computed returns: An application to the size related effect," *Journal of Financial Economics*, **12**:387-404.
- Brown, P., A.W. Kleidon and T.A. Marsh (1983), "New evidence on the nature of size-related anomalies in stock prices," *Journal of Financial Economics*, **12**:33-56.
- Chan, K.C., N. Chen and D. Hsieh (1985), "An exploratory investigation of the firm size effect," *Journal of Finance*, **46**:1467-1484.

- Chan, K.C., and N. Chen (1988), "An Unconditional Asset-Pricing Test and the Role of Firm Size as an Instrumental Variable for Risk," *Journal of Finance*, **46**:1467-1484.
- Chan, K.C. and N. Chen, (1991), "Structural and return characteristics of small and large firms" *Journal of Finance*, **46**:1467-1484.
- Chen, N., (1988), "Equilibrium asset pricing models and the firm size effect," in Elroy Dimson ed.: *Stock Market Anomalies*, Cambridge University Press, Cambridge, U.K.
- Clarkson, P. and J. Guedes and R. Thompson (1993), "Is Estimation Risk Priced? Is It Observable?" Cox School of Business, S.M.U., working paper.
- Denton, F. (1985), "Data Mining As An Industry," *Review of Economics and Statistics*, **February**:124-127.
- Dimson E. ed. (1988) *Stock Market Anomalies*, Cambridge University Press, Cambridge, U.K.
- Fama, E.F. (1991), "Efficient Capital Markets: II," *Journal of Finance*, **46**:1575-1618.
- Fama, E.F., and K.R. French (1992), "The Cross-Section of Expected Stock Returns," *Journal of Finance*, **47**:427-466
- Fama, E.F., and K.R. French (1993), "Common risk factors in the returns on stocks and bonds," *Journal of Financial Economics*, **33**:3-56
- Foster, F.D. and T. Smith (1992), "Assessing Goodness-of-Fit of Asset Pricing Models: The Distribution of the Maximal R^2 ," Working Paper, Fuqua School of Business, Duke University.
- Gibbons, M.R., S.A. Ross and J. Shanken (1989), "A Test of the Efficiency of a Given Portfolio," *Econometrica*, **57**:1121-1152.
- Jagannathan, R. and S. Viswanathan, (1988), "Linear Factor Pricing, Term Structure of Interest Rates and the Small Firm Anomaly," University of Minnesota working paper.
- Jegadeesh, N (1992), "Does Market Risk Really Explain the Size Effect," *Journal of Financial and Quantitative Analysis*, **27**:337-352.

- Keim, D.B., (1988), "Stock market regularities: A synthesis of the evidence and explanations" in Elroy Dimson ed.: *Stock Market Anomalies*, Cambridge
- Kothari S.P., J. Shanken and R.G. Sloan (1993), "Another Look at the Cross-Section of Expected Stock Returns," *Financial Research and Policy Studies Working Paper*, University of Rochester.
- Lo, A.W. and A.C. MacKinlay (1990), "Data-Snooping Biases in Tests of Financial Asset Pricing Models," *Review of Financial Studies*, **3**:431-468.
- Lo, A.W. and A.C. MacKinlay (1992), "Maximizing Predictability in the Stock and Bond Markets," MIT Working Paper.
- Roll, R., (1981), "A Possible Explanation of the Small Firm Effect," *Journal of Finance*, **36**:879-88.
- Schultz, P., (1983), "Transaction costs and the small firm effect: A comment," *Journal of Financial Economics*, **12**:81-88.
- Stoll H., and R. Whaley(1983), "Transaction costs and the small firm effect," *Journal of Financial Economics*, **12**:57-79.
- Schwert G.W. (1983), "Size and stock returns, and other empirical regularities," *Journal of Financial Economics*, **12**:3-12.