Is Size Dead?

A Review of the Size Effect in Equity Returns

Mathijs A. van Dijk[#]
Ohio State University and RSM Erasmus University
van-dijk_2@cob.osu.edu

Preliminary and incomplete – Comments solicited
January 2006

Abstract

This paper reviews 25 years of research on the size effect in international equity returns. Since Banz's (1981) original study, numerous papers have appeared on the empirical regularity that small firms have higher risk-adjusted stock returns than large firms. A quarter of a century after its discovery, the outlook for the size effect seems bleak. Yet, empirical asset pricing models that incorporate a factor portfolio mimicking underlying economic risks proxied by firm size are increasingly used by both academics and practitioners. Applications range from event studies and mutual fund performance measurement to computing the cost of equity capital. The aim of this paper is to review the literature on the size effect and synthesize the extensive debate on the validity and persistence of the size effect as an empirical phenomenon as well as the theoretical explanations for the effect. We discuss the implications for academic research and corporate finance and suggest a number of avenues for further research.

JEL Subject Codes

G12; G15; G31

Keywords

Size effect; international equity returns; CAPM; anomalies

[#] Visiting Scholar, Department of Finance, The Ohio State University, 736 Fisher Hall, 2100 Neil Avenue, Columbus, OH 43210, Phone (614) 292-1747, Fax (614) 292-2418. This paper has benefited from discussions with Karl Diether, Rüdiger Fahlenbrach, Kewei Hou, Pieter van Oijen, and Christian Wulff. Financial support from PricewaterhouseCoopers is gratefully acknowledged. All errors, misrepresentations, and omissions are my own.

1. Introduction

It has been 25 years since the first article appeared documenting that small stocks earn higher returns than traditional asset pricing models predict. Based on a study of U.S. equity returns over the period 1936-1975, Banz (1981) reported that small firms have considerably higher risk-adjusted returns than large firms and dubbed this finding "the size effect." Banz did not have a strong stance on the origin of the size effect and concluded that "It is not known whether size *per se* is responsible for the effect or whether size is just a proxy for one or more true unknown factors correlated with size" (p. 3, his emphasis). Since the early 1980s, a large body of research has been developed to address this issue and others related to it. This paper presents a review of the academic literature on the size effect in international equity returns.

The existence of a size effect in stock returns would have important implications for both practitioners and academics. If the higher returns on small stocks are due to a larger exposure to an underlying risk factor not incorporated in standard asset pricing models, firms should compute their cost of equity capital on the basis of an asset pricing model that takes this source of risk into account. The issue whether small stocks yield higher returns than large stocks and whether this is a compensation for risk is surely of interest from the perspective of asset management. Moreover, a risk-based explanation for the size effect would not only change the academic view on the validity of alternative asset pricing models, but would also have an impact on research methodologies such as event studies and mutual fund performance evaluation.

The early research on the size effect concentrates on the U.S. stock market. Our review starts with a survey of the empirical studies up to and including Fama and French (1992). The appearance of this paper marks a turning point in the literature. Fama and French provide a synthesis of the anomalies identified in previous studies and underline the flaws of the Capital Asset Pricing Model (CAPM). Subsequent research focuses on explaining the size effect. One possible explanation is that the empirical evidence is the result of data snooping. This necessitates out-of-sample tests. In order to evaluate efforts to address the data snooping concern, we review the empirical evidence on the size effect in international equity markets. Naturally, the study of the size-return relation in different countries is also interesting in its own right. A second possibility is that there are important methodological shortcomings in the empirical tests that uncovered the size effect. We dedicate a substantial part of this review to statistical and methodological critique on existing empirical studies. Third, since the early 1990s a passionate academic debate has evolved on the issue whether systematic risk factors can account for the size effect. Providing a synthesis of this debate is one of the main goals of this paper. The final part of our survey assesses the implications of the academic research on the size effect for practitioners as well as academics and offers suggestions for future research.

In contrast to previous review papers on the size effect that merely present a concise summary of the available empirical evidence (e.g. Hawawini and Keim (1995, 2000)), we present an overview of the methodological and statistical issues that play a role in the empirical literature. Moreover, earlier survey papers do not evaluate potential explanations for the size effect, while this paper assesses the validity of the empirical evidence as well as the merits of competing explanations for the size effect. In recent years, we have witnessed a surge in the application of empirical asset pricing models that include a risk factor constructed to capture the size effect. Both academics and practitioners increasingly implement the Fama and French (1993) three-factor model (or a variation thereof) for a wide variety of purposes. The apparent tension between the ongoing academic debate and the widespread use of these models warrants a critical review of the available evidence and the theoretical justifications of the size effect.

The remainder of the paper is structured as follows. Section 2 discusses the early empirical evidence on the size effect in the U.S. equity market. An overview of the international evidence on the size effect is presented in section 3. Section 4 examines various methodological objections raised as a reaction to the empirical studies on the size effect. In section 5, we critically review various explanations offered for the size effect. We assess the current state of the empirical and theoretical literature and discuss implications and directions for further research in section 6. Section 7 concludes.

2. Empirical evidence on the size effect for U.S. stocks

We begin our survey with an overview of the evidence on the size effect in U.S. stock returns that emerged in the 1980s and early 1990s. A discussion of later empirical research on the size effect in the U.S. is included in sections 4 and 5. Before we turn to an examination of the U.S. evidence, we present a concise description of the main methodologies used in the empirical literature.

2.1 Methodology

Three dominant methodologies can be distinguished in the vast empirical literature on the relation between firm size and returns. First, numerous papers use univariate sorting procedures. At time *t* all stocks are sorted into portfolios on the basis of their market capitalization.[1] Subsequently, the returns on the portfolios are calculated until the portfolios are rebalanced at time $t + 1$. Most studies compute equally-weighted returns and rebalance every year. The difference between the average return on the smallest and largest portfolio over the sample period is a measure for the size effect. Some studies compute risk-adjusted portfolio returns by estimating the beta of the portfolio. The method of Dimson (1979) is often used to adjust for nonsynchronous trading.

Starting with Fama and French (1992), many papers employ multivariate sorts on size and beta, as not only returns, but also betas tend to decline across size-sorted portfolios. Within each size portfolio, Fama and French sort stocks into portfolios on the basis of their beta over 2-5 years prior to time *t*. Subsequently, equally-weighted returns and portfolio betas can be estimated for each size-beta sorted portfolio. The main advantage of multivariate sorts is that the relationship between beta and returns can be examined independently of the size effect.

In order to investigate the asset pricing implications of the negative relation between size and returns, a cross-sectional test is required. A widely used approach is the methodology of Fama and MacBeth (1973). In each period, the cross-section of stock returns is regressed on beta and (the logarithm of) market value (and other variables, such as book-to-market equity). This allows for an explicit test of the hypothesis that beta and size explain the cross-section of stocks returns by computing time-series average of the coefficient on beta and size. The Fama-MacBeth methodology (or a variation thereof) is applied by the majority of studies on the size effect in the U.S.[2] In more recent studies, the time-series and cross-sectional data are often pooled and a Generalized Least Squares (GLS) estimation technique is applied.

---

[1] A number of studies use NYSE breakpoints for their sorting procedures. This implies that the number of stocks in the smallest size portfolio is relatively large, as AMEX/Nasdaq stocks generally have lower market values than NYSE stocks.

[2] Several studies use portfolios of stocks (formed on size and/or beta) in the cross-sectional regressions in order to reduce the “errors-in-variables” problem that arises when betas are estimated for individual stocks. However, when the formation of portfolios is not done carefully, this may lead to potentially large biases in the results. More discussion of this issue is presented in section 4.

2.2 Initial U.S. evidence on the size effect

This section summarizes the evidence on the size effect in U.S. stock returns presented by studies in the 1980s and early 1990s. An overview of the results of these studies is presented in Table 1. Banz (1981) is the first known empirical paper that presents evidence of a size effect. He analyzes all common stocks listed at the NYSE in the period 1936-1975. Banz reports that the smallest 20% of the firms earn a risk-adjusted return that is 0.4% per month higher than the remaining firms.[3] He runs a Fama-MacBeth regression of the returns on 25 size-beta portfolios on beta and market value. The coefficient on market value is negative and significant, indicating that small firms have higher risk-adjusted returns than large firms. This finding is robust to changes in the definition of the market portfolio. The size effect varies considerably over time, however, and assumes a negative (though insignificant) value in the period 1946-1955. Moreover, the size effect is not linear and is most pronounced for the smallest firms in the sample. Banz asserts that "… the size effect exists, but it is not at all clear why it exists. Until we find an answer, it should be interpreted with caution." (p. 17) Banz conjectures that higher uncertainty as a result of insufficient information about small companies may cause the size effect. Interestingly, this argument is closely related to the investor recognition hypothesis developed by Merton (1987).

Reinganum (1981) analyzes the size effect in a shorter, but broader sample of 566 NYSE and AMEX firms over the period 1975-1977. He finds that the smallest 10% of the firms outperform the largest 10% by 1.6% per month. The smallest of the 10 size portfolios has a beta roughly equal to 1 and a return of about 1 percent on a monthly basis in excess of the return on the equally-weighted market index. The largest size portfolio has a beta of 0.83 and underperforms the market by roughly 0.6% per month. Brown, Kleidon, and Marsh (1983) re-examine the size effect using the Reinganum data set of 566 firms over a longer sample period: 1967-1979. They find that there is an approximately linear relation between the average daily return on 10 size-based portfolios and the logarithm of the mean size of all firms in the portfolio. They also show that the size effect is unstable over time and is reversed in the period 1967-1975. Keim (1983) reports an average excess return of small stocks of 2.4% per month in a sample of NYSE and AMEX firms over the period 1963-1979. Again, the estimated size premium varies

[3] Unless explicitly stated otherwise, the size premia mentioned in this paper are not adjusted for risk. Risk-adjusted size premia are presented where available.

considerably over the years. Several measures of beta cannot explain the return difference between small and large firms.

Using a much broader sample of Nasdaq stocks than earlier studies, Lamoureux and Sanger (1989) find a risk-adjusted size premium of no less than 1.9% per month over the period 1973-1985. For NYSE/AMEX data, the premium is 1.2% per month. Lamoureux and Sanger note that the average firm size in a portfolio exhibits a positive monotonic relation with the average share price and a negative monotonic relation with the bid-ask spread. The authors also detect a strong January effect in both Nasdaq and NYSE/AMEX data. Their findings on the bid-ask spread and the January effect will be discussed in more detail in section 5.

Notwithstanding various important contributions in the decade after the original work by Banz, the literature on the size effect only really took off after the appearance of Fama and French (1992). Their paper synthesizes the size and book-to-market ("value") anomalies detected by earlier studies and demonstrates that the empirical shortcomings of the CAPM are too important to be ignored. Fama and French find that the smallest size decile outperforms the largest by 0.74% per month. Subsequently, they subdivide each size decile into 10 beta-sorted portfolios. Within each size decile, there is no relation between beta and return, but both average returns and post-ranking betas strongly decrease with firm size. This suggests that "… variation in beta that is related to size is positively related to average return, but variation in beta unrelated to size is not compensated." (p. 433) The results of Fama-MacBeth regressions confirm that while beta does not help to explain the cross-section of returns, size as well as book-to-market equity have significant explanatory power. The essentially flat relation between beta and returns has become known as the "beta is dead" conjecture. Subsequent research focuses on explaining the apparent breakdown of the CAPM and the causes of the size and book-to-market effects.

2.3 Conclusions

The early empirical literature based on U.S. data documents a return differential between small and large stocks that amounts to 0.4% to 2.4% per month (unadjusted for risk). Fama and French (1992) show that – within portfolios sorted on beta – small stocks have returns that are roughly 0.7% per month higher than large stocks in the U.S. Not only does the evidence suggests that the size effect survives correction for market risk, firm size also helps to explain the cross-section of stock returns. The size effect is strongest for very small stocks and there is evidence suggesting that stock returns are approximately linear in size as measured by logarithm of market value. Despite these strong empirical findings, several of the early studies note that the size effect exhibits considerable fluctuations over time and is reversed during some periods.

3. International evidence

An examination of the size effect in international equity returns is interesting for several reasons. First, a thorough understanding of the size effect in various countries facilitates the evaluation of corporate finance and investment decisions in those countries. Second, the strength of the size effect may depend on market characteristics such as the trading mechanism, the type of investors, and market efficiency in general. Third, the finding that the size effect exists in different markets and in different time periods would constitute a strong argument against data snooping concerns. This section presents a concise survey of 23 empirical studies of the size effect in non-U.S. markets. As these studies employ a wide variety of data selection techniques, sample sizes, and methodologies, we dedicate a subsection to gauging the reliability of their findings.

3.1 International evidence on the size effect

Since the late 1980s, a large number of studies have examined the magnitude of the size effect in an international context. Table 1 depicts an overview of the results of three empirical studies on the size effect in the U.K. Table 2 presents the (most recent available) estimates of the monthly size premium for 18 other individual countries and two groups of countries (emerging markets and Europe). The results of all studies presented in the tables are based on univariate sorting procedures on the basis of the market value of individual stocks.[4] The tables present the sample period and the number of securities studied, the number of size portfolios into which the securities are sorted, the average market value of the firms in the largest size portfolio relative to the average market value of the firms in the smallest size portfolio, and the average monthly return and beta estimate of the firms in the largest and the smallest size portfolio.[5]

---

[4] The exceptions are Dimson and Marsh (1999), who use index data for the U.K. and the U.S., and Elfakhani, Lockwood, and Zaher (1998), who use a multivariate sort on size and beta for Canada. For Canada, the information in Table 2 is based on averages across the beta-sorted portfolios.

[5] Most reported betas are adjusted for non-synchronous trading using either the Scholes and Williams (1977) method or the Dimson (1979) method.

Table 2 suggests that the international evidence on the size premium is remarkably consistent. Small firms outperform large firms (on a risk-unadjusted basis) in 17 of the 18 countries investigated as well as in a sample of emerging markets and in Europe. Monthly excess returns range from 0.13% for the Netherlands to 5.06% for Australia.[6] In 13 out of 18 countries, the excess return of small stocks over large stocks lies in the range of 0.4% to 1.2% per month. The first three rows of Table 1 reflect similar findings for the U.K. Hence, reported size effects outside the U.S. are substantial. This seems to indicate that the importance of data snooping issues, as expressed by e.g. Lo and McKinlay (1990) and Black (1993) and extensively discussed in section 5, is limited. The size effect, which was originally detected in U.S. data, is confirmed in numerous independent studies employing different data sets.

3.2 Methodology and robustness

The findings presented in section 3.1 seem to reinforce early U.S. evidence on the size effect. However, there are a number of important caveats. First, it is hard to judge whether small firms also outperform large firms on a risk-adjusted basis. U.S. studies show that small firms generally have higher betas than large firms. Roughly half of the studies included in Table 2 report a measure of the systematic risk of the size-sorted portfolios, while the other studies do not attempt to estimate the magnitude of the size effect on a risk-adjusted basis at all. Important differences exist in the betas of small and large stocks. Surprisingly, about half of the studies that report betas indicate that they are higher for large stocks than for small stocks. Almost none of the international studies performs a cross-sectional test in order to investigate whether firm size can explain the cross-sectional variation in stock returns. This makes it hard to judge whether firm size is a priced risk factor in international equity returns.[7]

Second, the sample composition of several studies evokes doubts about the reliability of the results. Papers that study five years of data or less (Korea, Mexico, Turkey), fewer than 100 securities (Finland, Ireland, Taiwan), or sort stocks into 3 portfolios or less (China, Mexico, Singapore) seem unlikely to provide a reliable representation of the magnitude of the size effect.

---

[6] The estimates of the size premium reported in Table 2 are based on size-sorted portfolio returns that are unadjusted for risk. However, for New Zealand and Spain the portfolio returns are calculated as abnormal returns relative to the CAPM.

[7] An exception is Chan, Hamao, and Lakonishok (1991), who show that the performance of the size variable in cross-sectional tests is highly dependent on model specification and time period.

In addition, some of the studies that have not been published in reputable academic journals give rise to concerns about possible inaccuracies in the sample selection and the application of the methodology. These concerns are reinforced when the reported size effects are extraordinarily large (5.1% per month for Australia; 4.2% for Mexico; 3.4% for Turkey).

Finally, few of the international studies perform a thorough analysis of the robustness of their results to, among other things, sample selection, return measurement interval, market index, extreme returns, and delisting bias. As outlined in section 4, checking whether the results are robust to dealing with these issues can be vital. Only a very limited number of analyses examines seasonality, which is an important issue in studies on the size effect (see the discussion in section 5).[8] There is some indication that – just as in the U.S. – the size premium is not robust in different time periods. Dimson and Marsh (1999) report evidence that the size premium reversed in the U.K. in the 1990s. The size premium was 5.9% over the period 1955-1988, while it amounted to –5.6% over the period 1989-1997. Dimson, Marsh, and Staunton (2002) employ large-cap and small-cap indices by Independence International Associates and FTSE International in order to get a crude indication of the sign and magnitude of the size effect in international equity markets. In 18 out of the 19 countries in their sample, the size effect appears to be reversed in the period after which an academic study on the size effect in that country appeared. They conclude that while there is a size *effect* (i.e., small firms perform differently from large firms), the question is "… whether we should continue to expect a size *premium* over the long haul." (p. 138) There is another robustness issue that is important in cross-country studies. Should the size of a firm be measured relative to the average size in its country? Annaert, Van Holle, Crombez, and Spinel (2002) report a significant size effect for 15 European countries. However, when they measure size relative to the average size of the firms within the same country, the size effect becomes insignificant. While the former approach makes it hard to distinguish the size effect in asset returns from a country effect, the latter approach ignores the fact that the largest firms from a small country may be relatively small in a European context. These firms should be expected to earn relatively high returns if the size effect holds and European markets are integrated.

[8] Dimson, Marsh, and Staunton (2002, chapter 9) report that the January effect is not present in U.K. stocks over the period 1955-2000.

3.3 Conclusions

The international evidence on the size effect is remarkably consistent at first sight. Small firms seem to outperform large firms in a large number of both developed and developing capital markets. However, the reliability of most international evidence is limited. Many studies employ small data samples, short estimation periods, and/or seem to suffer from data problems. Few studies present a comprehensive analysis of risk-adjusted returns and almost none of the studies use a formal cross-sectional asset pricing test of the size effect. The small number of studies that provide a thorough investigation of the robustness of the results indicate that the size effect was reversed in later periods for many countries. A more thorough analysis of the size effect in international equity returns is needed in order to defy the data snooping argument.

4. Methodological critique

Since the first studies on the size effect appeared in the early 1980s, a host of articles have been published on the validity of the methodologies employed to uncover the size effect. Some of this research primarily deals with the (sorting) methodologies used to evaluate the size affect, while others studies criticize empirical asset pricing tests in a more general context. This section assesses the most prominent methodological criticisms and their bearing on the reliability of the empirical evidence on the size effect.

4.1 The pitfalls of sorting methodologies

Berk (2000) criticizes the Fama and French (1992) methodology of sorting stocks into size portfolios first and then sorting stocks within each size portfolio into portfolios based on beta. The goal of this multivariate sorting technique is to investigate whether the CAPM has explanatory power within each size decile. Berk shows that this technique is biased toward rejecting whatever asset pricing model is examined in the second sorting step. The intuition is simple. By picking a variable in the first step that is empirically known to have a relation to stock returns, the return variation across groups is relatively large. This implies that the variation within groups is small and thus the statistical power to reject the null of a flat beta-return relationship is low. In a more general framework, Lo and McKinlay (1990) show that asset pricing tests may lead to misleading conclusions when properties of the data are used to construct the test statistics.

In general, data snooping (or data mining) biases emerge when a large number of studies analyze the same data set. In particular, the use of size-sorted portfolios may be affected by these data snooping biases. Lo and McKinlay examine to what extent classical statistical inference is influenced by the use of attribute-sorted portfolios in the empirical asset pricing literature. While sorting stocks into portfolios reduces the measurement error and enhances the power of the tests, grouping securities by some characteristic that is empirically motivated may lead to incorrect rejections of the null-hypothesis that the asset pricing model is true. The intuition is as follows. In an asset pricing test, one is interested in the magnitude of the abnormal returns (alphas). Tests based on combinations of alphas for portfolios of securities may be more powerful. However, estimated alphas can be regarded as the sum of true alphas and a measurement alphas. If researchers base the choice of the characteristic on which the securities are grouped on an empirical analysis of the (same) data only, there is no way of knowing whether any resulting cross-sectional relation between the alpha of a portfolio and the characteristic is due to a relation between the characteristic and the true alphas or a relation between the characteristic and the measurement error. Lo and McKinlay show that the true size may be up to 100% when the size of the traditional statistical test used is 5%. This does not necessarily imply that the size effect is spurious, as there may be a relation between firm size and true alphas. Statistical tests should take account of the bias described above and should preferably be performed on different data. However, "Much more convincingly would be the empirical significance of size (…) that is based on a model of economic equilibrium in which the characteristic is related to the behavior of asset returns endogenously." (p. 465). Accounting for previous analyses of the same data set in the statistical tests is extraordinarily complex. Therefore, analyzing many different data sets and time periods seems to be the most straightforward way to cope with the data snooping critique. More discussion of the data snooping argument is provided in section 5.4.1.

4.2 Size picks up any omitted risk factor

Although variables related to a firm's stock price can be used to detect flaws in asset pricing tests, it is not surprising that they contain information about the cross-section of expected returns. The argument is formalized by Berk (1995). He shows that if the asset pricing model employed is incorrect (or if the empirical specification is incorrect), firm size (measured by market capitalization) will always be inversely related with the part of return not explained by the model.

The intuition is clear. Of two firms with the same size (in the sense that end-of-period cash flows are exactly equal), the firm with riskier cash flows has a lower market value and, by definition, a higher expected return. Therefore, market value tends to be negatively correlated with all priced risk factors. Hence, if there is an omitted risk factor, market value will show up significantly in cross-sectional tests. Consequently, a significant coefficient on size in Fama-MacBeth regressions does suggest that the asset pricing factor is incorrect or misspecified, but size is not necessarily a proxy for the omitted risk factors in the model. Moreover, given the large measurement problems (as discussed by e.g. Fama and French (1992)), Berk argues that the conclusion that the CAPM is incorrect is too strong on the basis of current evidence. This argumentation does not necessarily imply that there is no priced risk factor related to size. In an unpublished manuscript, Berk (1996) investigates whether the reported evidence of the size effect can be attributed to a relation between returns and firm size (as measured by non-market related variables). He analyzes the relationship between the book value of assets, the book value of undepreciated property, plant, and equipment, annual sales, and the number of employees vis-à-vis the monthly returns of all NYSE firms over the period 1966-1987. Both multivariate sorts and Fama-MacBeth regressions show a negative relation between returns and market value, but no relation between returns and firm size. This indicates that the cross-sectional explanatory power of market value does not stem from a relation between firm size and returns.[9]

4.3 Mismeasurement of the market portfolio

The CAPM implies both that the market portfolio (defined as the market-weighted average of all assets) is mean-variance efficient and that there is a linear relation between expected returns and betas. However, as the market portfolio is unobservable, empirical tests of the CAPM are inevitably flawed. This argument forms the basis of the critique by Roll (1977) and Ross (1977). Roll argues that a correct test requires the inclusion of every individual asset in the market portfolio, which is not feasible. The evidence put forward by Stambaugh (1982) suggests that this problem is less severe than it seems, as inferences about the CAPM are similar for different market portfolios including bonds, real estate, and consumer durables in addition to common stocks. However, Roll and Ross (1994) show that OLS estimates of the cross-sectional relation

[9] Interestingly, Brown, Kleidon, and Marsh (1983) report a strong correlation (0.80) between the market capitalization and the book value of total assets for the firms in their sample, suggesting that the results are not sensitive to the measure of size employed, contrary to Berk's (1996) finding.

are very sensitive to the choice of the index and even indices close to the mean-variance efficient frontier can produce zero slopes. Sampling error exacerbates these problems.[10, 11]

Kandel and Stambaugh (1995) demonstrate that OLS estimates of the slope of beta in a cross-sectional regression of expected stock returns can be arbitrarily close to zero when the market index used is arbitrarily close to the mean-variance efficient frontier. On the other hand, a near perfect linear relation between beta and expected returns can be observed if the market index employed is far from efficient. The analysis of Kandel and Stambaugh indicates that the use of Generalized Least Squares (GLS) instead of Ordinary Least Squares (OLS) can considerably reduce this problem. The GLS estimate of the slope is positive as long as the expected return on the index proxy exceeds the expected return of the minimum variance portfolio. Note that GLS assumes that all (covariance) parameters are known. The use of feasible GLS in empirical studies introduces further sampling error that may affect the results.

Measuring the market portfolio is an intricate problem in both asset pricing tests and practical applications. A clear conclusion of the literature on this issue is that GLS estimation should be used in empirical work, which is not yet common practice. It is less obvious what the implications are for the implementation of asset pricing models like the CAPM for, e.g., calculating a firm's cost of equity capital. While virtually all applications currently rely on a market proxy based on common stocks only, it is not unimaginable that broader market portfolios that include other asset classes will be used more often in the future. For example, business valuation consultancies can offer cost of capital estimates for individual firms and industries that are based on alternative (broader) market proxies. Therefore, providing further insights into the impact of the market portfolio choice continues to be a valuable research objective.

---

[10] Black (1993) makes the more general point that if researchers use a market portfolio that differs from the true market portfolio, betas will be estimated with error. It is not unlikely that stocks that seem to have low betas will on average have higher betas when the true market portfolio would be used. This invalidates the analysis of whether the higher return on small stocks constitutes a premium for extra market risk. This problem has a bearing on the Fama-MacBeth methodology for detecting priced risk factors.

[11] Ferguson and Shockley (2003) demonstrate that if the market proxy used in empirical tests of the CAPM is equity-only, characteristics correlated with a firm's relative leverage and relative distress (such as firm size) will appear to explain returns. A three-factor model including the market and portfolios formed on the basis of leverage and distress outperforms the Fama and French (1993) three-factor model in cross-sectional tests, but not in time-series tests. Moreover, the high return on a portfolio of firms with low leverage and a low Z-score is not in line with the theoretical predictions.

4.4 Risk-loadings are time-varying

The CAPM is a static model and empirical tests often assume that betas are constant over time. However, in real life the relative risk of a firm's cash flows is likely to fluctuate over time and depend on the business cycle. Conditional versions of the CAPM take this variability into account by making expected returns conditional on the information available to investors at a given point in time. Jagannathan and Wang (1996) examine the ability of a conditional version of the CAPM to explain the cross-sectional variation in returns of 100 size-beta sorted portfolios of NYSE and AMEX stocks over the period 1962-1990. Using Fama-Macbeth regressions, the authors show that a conditional CAPM is able to explain roughly 30% of the cross-sectional return variation, compared to only 1% for the static CAPM. A second contribution of Jagannathan and Wang's paper is the inclusion of a measure of human capital in the market portfolio, in response to the market proxy problem discussed in section 4.3. This specification of the conditional CAPM explains roughly 50% of the cross-sectional variation in average returns and leaves no additional explanatory power for firm size.

In recent years, various studies have appeared that provide further evidence on the ability of conditional asset pricing models to explain the cross-section of stock returns. Lettau and Ludvigson (2001) analyze whether a conditional version of the Consumption CAPM (CCAPM) captures cross-sectional variation in the returns of 25 portfolios of U.S. stocks sorted on size and book-to-market. They find that this model performs much better than unconditional specifications of the CAPM as well as the CCAPM and about as well as the Fama-French three-factor model. Furthermore, the conditional CCAPM eliminates the residual size effect in the CAPM. Santos and Veronesi (2005) test a conditional CAPM in which labor income is the main state variable. They report no evidence of a size effect. Lewellen and Nagel (2005) question whether the conditional CAPM can explain asset pricing anomalies. An advantage of their approach is that they estimate betas over short windows and do not need to rely on a proxy for investors' information sets. In addition, they focus on the model's ability to explain time-series (instead of cross-sectional) variation in the returns of 25 portfolios sorted on size and book-to-market over the period 1964-2001. While betas vary considerably over time, they do not vary enough to explain known anomalies. Lewellen and Nagel find no evidence that portfolios of small stocks (un)conditionally outperform portfolios of large stocks in their sample.

Conditional asset pricing models are intuitively appealing, but whether they can explain asset pricing anomalies is subject to debate. There appears little evidence of a size effect in conditional specifications of the CAPM. What does this imply for implementations of asset pricing models in academic research and in practice? It seems unlikely that conditional versions of the CAPM will soon be used in event studies and cost of capital estimations, so the performance of unconditional models remains a relevant research topic. Lettau and Ludvigson (2001, p. 1281) suggest that the evidence indicates that the Fama-French factors are "… mimicking portfolios for risk factors associated with time variation in risk premia." Section 5 discusses the current body of evidence on this interpretation.

4.5 Return measurement interval

The CAPM does not provide guidance on the choice of horizon. The common choice of monthly returns is generally based on data considerations. Using relatively long horizons may prevent problems related to non-synchronous trading and seasonal patterns in returns. Brown, Kleidon, and Marsh (1983) find that the use of monthly instead of daily data does not markedly affect the results. Handa, Kothari, and Wasley (1989) show that betas of high-risk securities increase and betas of low-risk securities decrease with the return interval. Both the cross-section of monthly and the cross-section of annual stock returns is better explained with annual betas than with monthly betas. Neither monthly betas nor firm size provides additional explanatory power in Fama-MacBeth regressions when betas are measured on the basis of an annual return interval.[12]

4.6 Conclusions

There are four main streams of criticism on the methods employed to uncover the size effect. First, several papers have pointed out that sorting methodologies can lead to incorrect rejections of the CAPM as (i) sorting on size and then on beta reduces variation in returns within size deciles and (ii) data snooping biases may arise when sorts are based on empirical justifications only. Hence, caution should be applied in interpreting the results of sorting techniques. The lack of theory to explain the size effect not only affects sorting techniques, but also makes the Fama-Macbeth methodology vulnerable to data snooping arguments. A second fundamental critique on

[12] Kothari, Shanken, and Sloan (1995) also find that betas based on annual returns explain the cross-section of stock returns, but they focus on the book-to-market effect and do not report evidence on the size effect.

research on the size effect is that firm size will pick up any omitted risk factor (or misspecification) in an asset pricing model, as riskier firms generally have lower market values. The only paper we are aware of that examines non-market based measures of size does not find a relation between size and returns for U.S. stocks. A third important issue that affects all empirical tests of asset pricing models is the market proxy problem. Measurement issues may have an important impact on the evidence for the size effect, but are rarely discussed in detail in empirical studies. Finally, most empirical asset pricing studies that report a size effect assume that the risk loadings in the model are constant over time. Testing conditional versions of asset pricing models, several studies show that the size effect in equity returns disappears. The issue whether factors incorporated in empirical asset pricing models to capture the economic risks that drive the size effect proxy for time variation in risk premia is addressed in the next section.

5. Explanations for the size effect

In the past decade, the size and book-to-market effects have become the subject of one of the most heated debates in financial economics. At the heart of this debate is the question why small firms earn higher returns than traditional asset pricing models predict. A number of papers contend that the systematic risk of a stock is driven by multiple risk factors, and firm size is a proxy for the exposure to one of these factors. In fact, the multifactor model of Fama and French (1993) – that includes a factor constructed to capture this source of risk – is increasingly used by practitioners and academics for a variety of purposes. This interpretation of the size effect has resulted in a great deal of controversy, and a substantial number of researchers make the case that the size effect is little more than a statistical fluke. An alternative interpretation is that the size premium is a compensation for trading costs or liquidity risk. While this was first put forward in the mid 1980s, we have gained a lot of additional knowledge of liquidity risk in equity returns in recent years. This section aims at providing a comprehensive overview of the academic debate on the causes of the size effect.

5.1 It's risk!

Fama and French's 1992 paper contends that if assets are priced rationally, their results suggest that the systematic risk of a stock is multidimensional. The empirical case for the systematic risk story behind the size and book-to-market effect is presented in Fama and French (1993). The authors construct mimicking portfolios for the underlying risk factors related to size and book-to-market and argue that these portfolios capture the variation in returns. The mimicking portfolios are constructed as follows. Firms are sorted into three book-to-market groups and two size groups and six portfolios are created from the intersections of these. Subsequently, a small minus big (SMB) portfolio is constructed by subtracting the average return on the three big-stock portfolios from the average return on the three small-stock portfolios. A similar procedure is used for the construction of the high minus low (HML) portfolio that mimics the risk factor underlying the book-to-market effect. Fama and French show that the market, SMB, and HML portfolios capture a substantial part of the time-series variation in the returns of 25 stock portfolios formed on size and book-to-market over the period 1963-1991. Fama and French take this as evidence that size and book-to-market indeed proxy for sensitivity to common risk factors in stock returns.

Fama and French (1992, 1993) leave the issue of how economic fundamentals produce the common variation in return that is picked up by the SMB factor for further research. In their 1995 follow-up paper, they express this as follows: “Size and BE/ME remain arbitrary indicator variables that, for unexplained economic reasons, are related to risk factors in returns.” (p. 131) In this paper, they show that size and book-to-market factors exist in earnings. First, they show that size and book-to-market equity are related to firm profitability. High book-to-market firms have low earnings for a long period around portfolio formation. The relation between firm size and profitability is largely due to the 1980s, though. Subsequently, they construct SMB and HML factors in earnings, which are able to explain a considerable part of the time-variation in earnings of the portfolios. Fama and French take this as a strong indication that common risk factors in fundamentals (earnings) drive the risk factors in returns. Third, they find that the SMB factor in returns is related to the SMB factor in fundamentals. Fama and French (1996) show that their three-factor model also captures the returns on portfolios formed on the basis of other anomalies, such as cash flow to price and the long-term return reversal documented by DeBondt and Thaler (1985). Fama and French argue that the empirical success of the three-factor model indicates that it is an equilibrium pricing model, a three-factor version of Merton’s (1973) Intertemporal CAPM (ICAPM) or Ross’s (1976) Arbitrage Pricing Theory (APT).

While this evidence is suggestive of a fundamental explanation for the size effect, Fama and French do not address the issue which state variables produce variation in earnings and returns related to size and book-to-market. Chan, Chen, and Hsieh (1985) argue that stock returns react to changes in the economic environment. For example, in an economic downturn the required risk premium may rise. This suggests that economic risk factors, such as default risk, may explain the size effect in equity returns. They use the Fama-MacBeth regression technique to examine whether the returns on 20 size-ranked portfolios are related to the market portfolio and a number of macroeconomic variables. The data set consists of all NYSE firms over the period 1953-1977. While the difference in raw returns between the smallest stocks and the largest stocks is equal to 11.5% per annum, the yearly risk-adjusted return difference is only 1.5%. Almost half of the difference in raw returns can be explained by the spread between low-grade and government bonds, which can be regarded as a measure of changing risk premia. This suggests that the macroeconomic variables essentially capture the size effect. In a related paper, Chan and Chen (1991) argue that the size effect can be explained by underlying systematic risk faced by small firms. Their study indicates that small firms are generally “marginal firms” or “fallen angels” that have lost market value due to bad performance and tend to be firms that are not run very efficiently and have high leverage. Chan and Chen construct two size-matched portfolios in order to measure this effect. The first portfolio consists of firms that have recently cut dividends and the second portfolio consists of firms with high leverage. These two variables absorb the role of firm size in the Fama-MacBeth regressions. Chan and Chen argue that the exposure to these types of risk by small firms is not likely to be captured by a market index heavily weighted toward large firms. Vassalou and Xing (2004) investigate the relation between the size and book-to-market effects and default risk, defined as the risk that a firm fails to service its debt obligations. The authors estimate the default likelihood for up to 4,200 U.S. firms over the period 1971-1999 on the basis of contingent claims theory. The size effect turns out to be only statistically significant within the highest default risk quintile. Vassalou and Xing show that while the SMB and HML factors contain some default-related information, default risk cannot account for the explanatory power of the Fama-French three-factor model.

Dichev (1998) further investigates the hypothesis formulated by Chan and Chen (1991) that the size effect may be explained by financial distress. Dichev specifies a distress factor by measuring the probability of bankruptcy. The finding that bankruptcy risk is systematic would support a distress risk explanation of the size effect. There is a well-developed literature on bankruptcy prediction that provides measures of ex ante bankruptcy risk. The author examines

NYSE, AMEX, and Nasdaq stocks over the period 1981-1995 and finds that bankruptcy risk is not associated with higher stock returns, which sheds doubts on a risk-based explanation of the size effect. Campbell, Hilscher, and Szilagyi (2005) show that U.S. firms with a high probability of bankruptcy have a high loading on the SMB factor. However, inconsistent with the conjecture that the size premium is a compensation for distress risk, these firms do not earn higher returns.

The interpretation that firm size proxies for a firm's exposure to an underlying risk factor is controversial. Ferson, Sarkissian, and Simin (1999) show that the evidence that attribute-sorted portfolios (such as SMB and HML) capture a substantial part of the time-series variation in stock returns does not provide fundamental additional insights into the issue whether a reported anomaly is related to a systematic risk factor for which the attribute-sorted portfolio proxies. Their argument is illustrated using the "alpha factor asset pricing model." In a simulation experiment, Ferson et al. show that constructing a portfolio mimicking an "alphabet factor" unrelated to risk yields results similar to Fama and French (1993). They conclude that even if an observed phenomenon is completely unrelated to systematic risk, attribute-sorted portfolios based on that phenomenon will appear to be useful risk factors. This underscores the importance of a sound theoretical justification for the size effect. Among others, Black (1993), MacKinlay (1995), and Ferson and Harvey (1999) warn for the data snooping peril in the absence of theory for the existence of the size effect. Black (1993) puts forward an alternative explanation for why low beta stocks may have higher expected returns than the CAPM predicts. He argues that borrowing restrictions, such as margin rules, bankruptcy laws that limit lender access to a borrower's future income, and tax rules that limit deductions for interest expenses, may cause investors interested in high risk portfolios to bid up the prices of high beta stocks. Without borrowing restrictions, investors could borrow money and invest the proceeds in the market (or in low beta stocks). MacKinlay (1995) shows that CAPM deviations are very difficult to detect empirically without a specific alternative economic theory. Ferson and Harvey (1999) find that predetermined economic variables (such as the dividend yield of the S&P 500 index and several bond yield spreads) that conditional asset pricing studies show to have some power in explaining the time-series of stock returns also provide significant cross-sectional explanatory power in addition to the Fama-French factors. This explanatory power does not disappear when possible time-variation in the exposures to the Fama-French factors is accounted for. This evidence suggest that the hypothesis that the Fama-French factors fully explain the cross-section of stock returns can be rejected. Daniel and Titman (1997) show that firms characteristics rather than factor loadings on the SMB and HML portfolios determine expected returns. Within portfolios formed on size, there is basically no relation between returns and loadings on the SMB factor. This constitutes evidence against a financial distress interpretation of the SMB factor.[13, 14]

5.2 It's liquidity!

The CAPM and other traditional asset pricing models abstract from the influence of liquidity and other market microstructure issues. Several authors have suggested that the size effect may simply be a result of higher trading costs for small stocks, for which investors have to be compensated. More recently, a number of asset pricing models have been put forward that include a priced risk factor related to liquidity risk.

The first paper to relate the size effect to liquidity is Stoll and Whaley (1983). This study investigates NYSE stocks sorted into 10 size portfolios over the period 1960-1979. They examine abnormal returns on small stocks after transaction costs and find that it is not possible to earn abnormal risk-adjusted returns after accounting for transaction costs (at horizons up to one year). Schultz (1983) extends the analysis of Stoll and Whaley (1983) and finds that transaction costs cannot explain the size effect for NYSE and AMEX data over the period 1962-1975. His findings indicate a significant (risk-adjusted) size effect of 2% per month after transaction costs for an investment horizon of one year. Schultz concludes that the size effect cannot be solely explained by differences in transaction costs between small and large firms.

Amihud and Mendelson (1986) present a theoretical model in which expected returns are increasing in the bid-ask spread. The basic intuition is that investors require a compensation for expected trading costs. The model predicts that investors with longer holding periods select securities with larger spreads. Therefore, the larger the spread, the smaller the compensation

[13] Heston, Rouwenhorst, and Wessels (1999) present similar results for a large sample of European stocks. Davis, Fama, and French (2000) indicate that over the extended sample period 1929-1997 the U.S. size effect is too small to accurately distinguish between the risk model and the characteristics model.

[14] An alternative risk-based explanation for the size effect is that small stocks exhibit higher idiosyncratic risk. Modern portfolio theory argues that idiosyncratic risk is irrelevant to investors, because it can be diversified away in a large portfolio. Malkiel and Xu (1997, 2004) argue, however, that idiosyncratic risk may be priced when not all investors (are able to) hold the market portfolio and because portfolio managers are often called upon to explain why they held on to specific stocks that declined in value. The authors extend the dataset of Fama and French (1992) to the year 2000 and show that incorporating a measure of idiosyncratic risk absorbs the size effect in Fama-MacBeth regressions. Other papers, e.g. Goyal and Santa-Clara (2003), confirm that stock returns are related to idiosyncratic risk, but do not look at the relation with the size effect.

required for an additional increase in the spread, as transaction costs are amortized over a longer holding period. Hence, the relation between expected return and spread is concave. This theoretical prediction is borne out in an analysis of the data set of Stoll and Whaley (1983). Fama-MacBeth regressions reveal that beta and spread exhibit significant cross-sectional explanatory power for the returns on 49 spread-beta portfolios. The size effect is not significant and may thus be a consequence of a "spread effect." Eleswarapu and Reinganum (1993) criticize the sample selection criteria of Amihud and Mendelson (1986). In particular, their sample excludes very small stocks, because they require a return history of eleven years. Investigating data that are updated through 1990 and include a much larger number of firms, they report that size is the only variable that shows up significantly in the Fama-Macbeth regressions. The coefficient on the spread variable is insignificant. The results indicate that the spread effect cannot fully explain the size effect.

Several other authors examine the relationship between expected stock returns and various measures of liquidity. Brennan and Subrahmanyam (1996) contend that the quoted bid-ask spread is a noisy measure of liquidity, as many transactions occur inside or outside the spread and the price impact of a trade is not taken into account properly. The authors study intraday transaction data for NYSE stocks over the period 1984-1988 and use market microstructure models to estimate fixed and variable transaction costs from the transactions data. Both fixed and variable transaction costs are positively and significantly related to returns, consistent with a significant risk premium on (il)liquidity. Datar, Naik, and Radcliffe (1998) use the turnover rate (number of shares traded as a fraction of shares outstanding) as a proxy for liquidity of NYSE stocks in the period 1962-1991. They find that individual stock turnover explains the cross-section of stock returns, even after controlling for firm size and book-to-market. Neither Brennan and Subrahmanyam (1996) nor Datar et al. (1998) explicitly examine the relationship between the size and liquidity factors. Amihud (2002) employs the Fama-MacBeth methodology to study the relation between returns of NYSE stocks over the period 1964-1997 and the ratio of absolute stock return to dollar trading volume, which is an easily obtainable measure of price impact. The variable is highly correlated with market value. The results support the hypothesis that illiquidity is priced. Both the illiquidity measure and firm size have a significant effect of the expected sign, however, suggesting that the illiquidity variable does not capture the size effect completely.

Recent studies consider the possibility that market liquidity is a priced state variable. If the returns of small stocks are more sensitive to this state variable, part of the size effect may be related to (il)liquidity risk. Amihud (2002) runs time-series regressions of the returns on 5 size-sorted portfolios on expected and unexpected market liquidity. Small firms are indeed more sensitive to market liquidity. Time-variation in the size effect may therefore be related to time-variation in the price of liquidity risk. The explanatory power of the regressions is low and alphas are large, however, suggesting that the liquidity effect captures only a minor part of the time-series variation in returns. Pastor and Stambaugh (2003) study whether market liquidity is priced in the returns of NYSE, AMEX, and Nasdaq over the period 1966-1999. In contrast to Amihud (2002), they do not perform a time-series analysis, but examine whether liquidity risk is a priced risk factor. Their liquidity measure is a market-wide average of the effect of trading volume on day t "signed" by the return on day t on the return on day $t + 1$. The results indicate that liquidity risk is important in asset pricing. Portfolios of firms with high liquidity betas have substantially higher returns relative to the Fama and French (1993) three-factor model than portfolios of firms with low betas. Interestingly, the highest liquidity betas are concentrated in the smallest firms. Pastor and Stambaugh contend, however, that the relation between liquidity risk and firm size is not necessarily clear-cut and do not investigate whether size is a significant determinant of expected returns after correcting for liquidity risk. Acharya and Pedersen (2005) present a simple and appealing equilibrium model in which the expected return on a stock depends on its expected liquidity and on the covariances of its own return and liquidity with the market return and liquidity. The model is tested using Amihud's (2002) measure of liquidity and data for NYSE and AMEX stocks over the period 1962-1999. Cross-sectional tests reveal that the model has a higher explanatory power than the CAPM and the liquidity risk premia are economically significant. Small stocks have lower average liquidity and higher exposures to the three liquidity risk factors. The authors show that incorporating the liquidity risk factors improves the fit for portfolios of small stocks, but do not examine whether liquidity risk absorbs the size effect. Further research is called for to analyze the relation between the size effect and liquidity risk in detail.

A potential shortcoming in the liquidity explanation for the size effect – and indeed in other explanations as well – is that it does not seem to explain the strong seasonality in the size effect. As reported by Keim (1983) and discussed in more detail below, the size effect in the U.S. is to a large extent due to the month of January. Future studies should establish whether transaction costs and liquidity risk can account for this. Preliminary investigations by e.g. Lamoureux and Sanger (1989) indicates that transaction costs show limited seasonal behavior and are thus not likely to completely account for the size effect.

5.3 It's investor behavior!

While a number of studies on the book-to-market effect (notably Lakonishok, Shleifer, and Vishny (1994)) have pointed at causes embedded in the behavior of investors, similar types of explanations for the size effect have remained virtually unexplored. Yet, it is not implausible that the overreaction interpretation of the value effect holds some water for the size effect. The argument is that firms with high book-to-market ratio's ("value firms") are typically firms that have shown poor performance in the past. If investors over-extrapolate past performance, this will lead to a stock price of value firms that is too low, resulting in higher returns when the overreaction is eventually corrected. Papers such as Chan and Chen (1991) indicate that small firms also tend to be firms that have done poorly in the past. This suggests that extrapolation could be a driving force of the size effect as well.

An alternative explanation based on investor behavior that has been offered for the value effect is that investors like growth stocks and dislike value stocks. It seems equally legitimate to argue that investors prefer large stocks over small stocks. The findings of Daniel and Titman (1997) that firm size and book-to-market determine expected returns as characteristics of stocks rather than proxies for risk can be interpreted in this light. Lakonishok, Shleifer, and Vishny (1992) argue that agency relationships have a bearing on portfolio selection by professional money managers. It can be argued that investments in small stocks are harder to justify to sponsors. Moreover, the size effect may originate from incomplete information about small firms. The model of Merton (1987) predicts that less well-known stocks of firms with smaller investor bases have higher expected returns. An empirical analysis of the influence of investor recognition on the size effect is offered by Hou and Moskowitz (2005), who propose the average delay with which a firm's stock price reacts to information as a broad measure for market frictions. Price delay has a significant impact on the cross-section of U.S. stock returns over the period 1963-2001 and captures a substantial part of the size effect. The authors show that the results are most consistent with frictions associated with investor recognition.

5.4 It's a fluke!

5.4.1 Data mining and robustness

Among others, Black (1993) and Lo and MacKinlay (1990) argue that many researchers have worked on the same data to uncover the size effect and other asset pricing anomalies. Only the most successful, unusual, and striking results are published. This makes it impossible to assess their statistical significance, which also depends on the number of attempts needed for discovering a certain effect. Out-of-sample tests are needed to counter the data mining argument. Section 3 argues that the international evidence for the size effect is not unambiguous. Several authors investigate the robustness of the results for the U.S. market. Banz (1981) shows that the size effect in the U.S. varies considerably over the period 1926-1975. Keim (1983) confirms this observation. Handa, Kothari, and Wasley (1989) report different estimates of the size effect in the U.S. in three subperiods of their sample period 1941-1982. The size effect is negative (though insignificant) in the period 1941-1954. Davis, Fama, and French (2000) report a statistically significant monthly return difference of 0.33% for portfolios of NYSE, AMEX, and Nasdaq stocks with a market value below and above the NYSE median over the period 1929-1997.

There is evidence indicating that the size effect has diminished since the publication of Banz's paper in 1981. Dichev (1998) reports that the size effect is virtually non-existent in the period 1980-1995. Eleswarapu and Reinganum (1993) and Amihud (2002) detect no size premium over the sample periods 1980-1990 and 1980-1997, respectively. Horowitz, Loughran, and Savin (2000) find that large firms have slightly higher annual returns than small firms in a set of NYSE, AMEX, and Nasdaq data over the period 1979-1995. Dimson and Marsh (1999) report that small stocks outperformed large stock by 4.1% per year over the period 1955-1983, but underperformed large stocks by 2.4% between 1983 and 1997. Reinganum (1999) contends that the variability in the size effect over time is predictable in the U.S. Notably, large firms outperform small firms in adverse economic conditions.

A reversal of the size effect in certain periods does not necessarily imply that small firms do not earn higher returns on average than their beta suggests.[15] Bonds also occasionally outperform stocks over prolonged periods of time, yet few economists would dispute the

[15] Indeed, a risk-based explanation of the size effect would suggest that small stocks underperform large stocks with some frequency.

proposition that on average stocks yield higher returns than bonds as a compensation for their higher systematic risk. However, if these reversals occur often and/or over extended time periods, this may shed doubts on the reliability of the empirical findings. Figure 1 plots the annual, market-weighted return differential between the smallest and the largest size quintile of all NYSE, AMEX, and Nasdaq over the period 1927-2004.[16] Note that these returns are not adjusted for (market) risk. The average return differential amounted to 7.0% per annum over this period. There is a lot of variation in the size effect over time. There are three years in which the return on small stocks was over 50 percentage points higher than the return on large stocks, while in many years the difference is modest. In 35 (out of 78) years, the size premium is negative. Averaged over the past 25 years, the return difference is only 1.4% per annum and especially in the periods 1946-1957 and 1980-2000 small stocks have clearly not outperformed large stocks. While the graph does suggest that evidence for a size effect has become weaker in the past decades, it is not straightforward to draw inferences about the validity of the size effect. Figure 2 displays the annual returns on the U.S. market portfolio (in excess of the risk-free rate) and the SMB and HML mimicking portfolios over the same period. The SMB portfolio depicts a similar return pattern as Figure 1. Although the returns on the market and the HML factor are more consistently positive over time, they also exhibit considerable variation over time. The graphs do not unambiguously show that the evidence on the size effect is not robust or that the effect has disappeared in recent years. Nevertheless, it is remarkable that no research to date has addressed the question whether structural or institutional changes can account for the decrease in the size premium since the early 1980s. Has the liquidity of small stocks improved? Have private investors become more sophisticated? Has the expansion of investment strategies aimed at exploiting the size anomaly eliminated the effect? Or has there never been a reliable size effect?

5.4.2 Delisting bias

Shumway and Warther (1999) investigate the implications of the delisting bias in Nasdaq data and demonstrate that the size effect disappears when the bias is taken into account. A significant fraction of the returns associated with delistings (due to a merger or acquisition, bankruptcy, liquidation, or migration to another exchange) is not recorded in CRSP. Shumway and Warther

[16] We thank Ken French for providing the returns on the size quintiles, the market portfolio, and the SMB and HML factors on his website: http://mba.tuck.dartmouth.edu/pages/faculty/ken.french/

collect over-the-counter data on delisting returns and conclude that researchers should correct for the delisting bias by using a delisting return of –55% for the delisted stocks with missing data. Shumway and Warther re-examine the evidence of Lamoureux and Sanger (1989) of a size effect on the Nasdaq by sorting all Nasdaq stocks into 20 portfolios. Over the period 1972-1995, the return differential between the smallest and largest portfolio amounts to 2.7% on a monthly basis, while the size premium decreases to 0.93% per month when the delisting correction is applied. The former result is statistically significant, while the latter is not. This result is robust in several subperiods. Shumway and Warther (p. 2361) conclude that "... there is no evidence that there ever was a size effect on Nasdaq."

5.4.3 Extreme returns

Knez and Ready (1997) show that the size effect is driven by the extreme 1% of the observations.[17] They re-examine the Fama and French (1992) data with a robust regression technique – least trimmed squares – which trims a proportion of the observations and fits the remaining observations using least squares.[18] When the extreme 1% of the observations are trimmed, the Fama-MacBeth regressions do not yield a significantly negative, but a positive coefficient on firm size. The authors stress that the extreme observations should not be considered as outliers. However, the analysis does suggest that most small firms actually underperform big firms. The size effect seems to be due to a tiny fraction of the small firms that do extremely well. The authors speculate that this is related to the "turtle eggs" effect: most small stocks do not perform well, but this is compensated by a small number of extremely successful firms. While this is an interesting suggestion, more research is needed to understand this supposition.

5.4.4 Seasonality

While seasonality in the size effect does not imply that the size effect is a statistical aberration, the strong January effect in the return differential between small and large stocks is an important phenomenon that is little understood. Keim (1983) shows that 50 percent of the size effect is due to excess returns in January (amounting to 15.0% annually on average), much of which stems

[17] A similar result is reported by Downs and Ingram (2000).

[18] Note that Fama and French (1992) do trim the 0.5% most extreme book-to-market observations in the Fama-MacBeth regressions, but they do not trim the extreme size observations.

from the first five trading days. Keim puts forward two possible explanations (without an attempt to testing them). First, the tax-loss selling hypothesis predicts that toward the end of the year investors sell stocks that have declined in price during the year in order to take advantage of tax benefits. This effect could be especially important for portfolios of small stocks, as these may be biased toward shares that have experienced large price declines. Second, considerable amounts of information become available in January which leads to increased uncertainty. This may also be relatively important for small firms, as information is more costly for those firms.

Several other studies confirm Keim's findings for other data sets. Brown, Kleidon, and Marsh (1983) find a strong January size effect in NYSE stock returns over the period 1967-1979. Lamoureux and Sanger (1989) observe a January effect in Nasdaq returns over the period 1972-1985. They show that the smallest size decile outperforms the largest by no less than 10.4% in January. Again, a large part is due to the first five trading days. Remarkably, they report an almost perfect positive monotonic relation between firm size and excess return for the months February-December. Lamoureux and Sanger demonstrate that transaction costs exhibit hardly any seasonal behavior and consequently cannot fully explain the size effect. Daniel and Titman (1997) show that the size effect for NYSE stocks in the period 1963-1993 is to a large extent due to a strong size effect in January. Finally, Dimson, Marsh, and Staunton (2002) show that the January effect in CRSP data was large over the period 1926-2000 and almost entirely due to the superior performance of small firms.

Consistent with the tax-loss selling hypothesis, Roll (1983) finds a negative relation between the returns around January 1 and the returns over the previous year for NYSE/AMEX stocks over 1962-1980. Brown, Keim, Kleidon, and Marsh (1983) argue that the tax-loss selling argument would predict a July seasonal in the return of small stocks in Australia, as the Australian tax year ends in June. In contrast with this conjecture, they show that the Australian size premium is fairly stable across the months in the period 1958-1981. Berges, McConnell, and Schlarbaum (1984) also conclude that the tax-loss selling hypothesis cannot completely explain the January effect and is thus unlikely to explain the size effect. They analyze a sample of Canadian companies and show that the January effect is only slightly more pronounced after the introduction of the capital gains tax in Canada in 1973. Kato and Schallheim (1985) find a strong January-size effect in Japan over the period 1952-1980. As the tax regime in Japan generates no reason for tax-loss selling in December, this suggests that this explanation is incomplete.

Eleswarapu and Reinganum (1993) document important seasonal components in the size effect as well as in the impact of the bid-ask spread on equity returns. Looking only at the January months in the sample, both spread and size are priced, while in the non-January months, neither is priced. The analysis is repeated for the subperiod 1980-1990. Over this period, the size effect is absent, except in January. In that month, a significantly positive relationship between spread and return is detected, while the effect is significantly negative in the non-January months.

Figure 3 depicts seasonal patterns in the market-weighted return differential between the smallest and the largest size quintile of all NYSE, AMEX, and Nasdaq over the period 1927-2004. The graph suggests that the size effect in the U.S. is almost entirely due to higher returns on small stocks in January. The return differential amounts to almost 6% in January and is close to zero in all other months. Closer inspection of the origin of this seasonal effect reveals that there is a strong January effect in the returns of the smallest size quintile, while the returns of the largest size quintile exhibit little seasonal variation. The issue how the January effect and the size effect are related is vital for understanding the size effect, but remains an open question. In addition, a comprehensive study of seasonality in the size effect in international equity returns is warranted.

5.5 Conclusions

The academic debate on the source of the size effect is unresolved. Some of the empirical evidence points in the direction of an explanation based on underlying economic risks. However, a theoretical model describing these economic sources of risk is still to be developed. And, perhaps surprisingly, there does not seem to be an empirical relation between bankruptcy risk and the size effect. The theoretical foundations of explanations based on liquidity are stronger. Nevertheless, the empirical evidence is mixed and most studies analyzing liquidity risk do not systematically examine the relation with the size effect. A challenge to both explanations is the strong January seasonal observed in the size effect. Finally, the robustness of the size effect does not appear strong enough to dismiss the data snooping argument and evidence suggests that the delisting bias in CRSP data accounts for the size effect in Nasdaq data.

6. Implications

In this section we provide an evaluation of the empirical, methodological, and theoretical research on the size effect discussed in this survey. Although many of the early empirical studies identify a significant and consistent size premium in U.S. equity returns, the overall evidence of a size effect in the U.S. is not overwhelming and several papers report that the effect has disappeared after 1980. While a risk-based explanation of the size effect would imply that large stocks occasionally outperform small stocks, the instability of the size effect reinforces the concern that data snooping biases may have played a role in the discovery of the effect. The standard response to this critique is performing out-of-sample tests. A large number of papers report results that are consistent with the existence of a size premium in many countries. However, systematic cross-sectional asset pricing tests as well as more elaborate robustness checks are required in order to make a truly compelling case for the existence of a size effect in international equity returns.

Both sorting procedures and cross-sectional tests are vulnerable to data snooping. Therefore, developing theoretical explanations is essential for our understanding of this phenomenon. This is underlined by the fact that firm size will pick up any omitted risk factor in asset pricing tests. As market value and firm risk are inversely related, size will appear to add explanatory power to asset pricing models that are either incorrect or misspecified. This has the important implication that the market value of a firm cannot necessarily be regarded as a proxy for an underlying risk factor that is related to the size of the firm. At least one (unpublished) paper shows that other measures of firm size do not explain the cross-section of stock returns.

What omitted factor does market value pick up in asset pricing tests? Fama and French (1992, 1993, 1995, 1996) argue that market value proxies for an underlying systematic risk factor related to firm size. This view is corroborated by empirical evidence that a portfolio constructed to mimic this risk factor adds to the variation in returns explained by the market portfolio and that the size effect is absorbed by risk factors related to firm distress and default risk. However, the size effect cannot be explained by bankruptcy risk. This appears to be inconsistent with the view that firm distress risk drives the differences in returns of small and large stocks. The explanation of the size effect based on financial distress still lacks backing by a formal economic theory.

A second potential explanation for the size effect is liquidity. A stock’s expected trading costs as well as its exposure to one or more liquidity risks are likely to affect its expected return.

These effects may well be particularly important for small stocks. Various models have been developed to capture the role of liquidity in explaining the cross-section of stock returns. The available evidence indicates that liquidity is an important factor in asset pricing. However, most studies do not explicitly examine whether the size effect can be explained by liquidity factors and the few studies that do present mixed evidence. How the size effect and liquidity interact is an important area for future research.

A weakness in both explanations for the size effect discussed above is their apparent inability to account for the strong January seasonal in the return differential between small and large stocks. While several studies discuss this January effect in relation to the size effect, no convincing explanation has been offered. Several papers show that the tax-loss selling is unlikely to fully account for the January seasonal. This literature is still developing and additional empirical research on the January effect in the size premium in international markets as well as further theoretical explanations are needed to understand this phenomenon.

There has been some useful work on a number of other possible explanations for the size effect. Conditional versions of the CAPM and the CCAPM leave little room for firm size in explaining the cross-section of stock returns. If and how factors that capture the size effect in unconditional empirical asset pricing models proxy for time variation in risk premia is not clear. Idiosyncratic risk may also account for the role of firm size in cross-sectional tests, but this strand of the literature is relatively undeveloped. Finally, the explanation for the size effect based the “turtle eggs” effect is still embryonic. To these potentially fruitful avenues for further research we would like to add the question whether investor behavior can account for the size effect. Theoretical and empirical research on investor overreaction and the investor recognition hypothesis in relation to the size effect could yield important insights into the causes of the effect.

Concluding, we assess the empirical evidence for the size effect to be consistent at first sight, but frail at closer inspection. More empirical research is needed to establish the robustness of the size effect in international equity markets. At the same time, there is little consensus about the origin of the size effect. Size may proxy for underlying economic risk factors, but this argument is not supported by theory. Trading costs and liquidity risk may play a role, but empirical research needs to concentrate more efforts on the interaction between size and liquidity. Several other potential explanations have been suggested, but are not yet backed by comprehensive studies. Any theoretical explanation should address the issue why the size effect is especially pronounced in January.

Where does that leave academics and practitioners? Despite the ongoing debate about the merits of the Fama-French three-factor model, the model is increasingly used in studies of corporate events, in mutual fund performance evaluation studies, and in consultancy reports on the estimation of a firm's cost of equity capital.[19, 20] Fama and French (2004) rationalize these applications by underscoring that (i) one wants to measure the premia investors require to hold stocks, whether these capture the exposure to unknown state variable risks or are the result of irrational investor behavior, and (ii) in event studies and performance evaluation research one wants to account for known patterns in returns, whether they are sample specific or permanent. While these arguments may be more appropriate for the book-to-market effect, we are hesitant to recommend the application of an empirically-inspired asset pricing model while ambiguity exists about the robustness as well as the causes of the size effect it incorporates. We recognize the rationale of using the Fama-French three-factor model in evaluating whether portfolio managers achieve higher returns than expected on the basis of common investment styles. And researchers identifying new anomalies, patterns in returns, or profitable trading strategies may want to bear out that these are not different manifestations of the size effect. But in particular for calculating a firm's cost of equity capital, the three-factor model seems barely less crude than using the average stock return of that firm (or its industry) in the past. On a final note, it is interesting to observe that Fama and French seem more convinced about the book-to-market effect than about the size effect. Their 2004 review of the CAPM focuses on the literature on the book-to-market effect and in a 2005 paper on the CAPM and the value premium, they mention that, studying U.S. stock returns over the period 1926-2004, there is "… little evidence against CAPM pricing as the explanation for the premium in the average returns on small stocks." (p. 12)

[19] See e.g. Ibbotson Associates' and Standard & Poor's Corporate Value Consulting Risk Premium Report.

[20] Survey evidence presented by e.g. Graham and Harvey (2001) and Brounen, de Jong, and Koedijk (2004) indicates that it is not yet common practice in corporations in Europe and the U.S. to take account of the size effect in valuation and capital budgeting issues.

7. Concluding remarks

In the introduction to a special issue of the *Journal of Financial Economics* on the size effect Schwert (1983, p. 3) provides the following definition: "… average returns to small firms' stocks are substantially higher than any known capital asset pricing model predicts." Schwert contends that the empirical evidence for the size effect in the U.S. is strong, notwithstanding several methodological issues that have been raised. At the same time, he is not optimistic that the cause for the size effect will be understood soon. Therefore, Schwert does not expect that the size effect will be taken into account in capital budgeting issues, nor in performance evaluation for investment portfolios. The effect is, however, likely to be incorporated in event study methodology in order to account for seasonality. Schwert concludes that "… to successfully explain the 'size effect', new theory must be developed that is consistent with rational maximizing behavior on the part of all actors in the model." (p. 10)

In the more than 20 years that have passed since this publication, a huge literature has developed on the size effect. The size effect has been investigated empirically for a large number of countries and numerous studies have attempted to provide an explanation for the puzzle. Schwert's prediction about event study methodology has been proven right, as many recent event studies employ the Fama-French three-factor model as a benchmark for expected returns. Contrary to Schwert's expectation, academic research evaluating the performance of hedge and mutual funds also relies heavily on multifactor models including a "size risk" factor. In addition, the Fama-French three-factor model is nowadays discussed by some corporate finance textbooks and consultancies publish reports containing estimates of the size effect.

The review presented in this paper identifies a number of important gaps in the theoretical and empirical literature on the size effect. The available international empirical evidence is feeble, as several important methodological and statistical issues are not addressed sufficiently and because the robustness of the effect has not been established convincingly. The theoretical explanations for the size effect that have been offered in the literature are incomplete. This renders the application of the Fama-French model by corporate financial managers premature. Size may not be "dead," but more empirical and theoretical research is required in order to revive its ailing condition. Schwert's assertion that new theory must be developed consistent with rational maximizing behavior is as valid today as it was in 1983.

References

Acharya, V.V., and L.H. Pedersen, 2005, "Asset Pricing with Liquidity Risk," *Journal of Financial Economics*, 77, 375–410.

Aksu, M.H., and T. Onder, 2003, The Size and Book-to-Market Effects and Their Role as Risk Proxies in the Istanbul Stock Exchange," *SSRN Working Paper*.

Amihud, Y., 2002, "Illiquidity and Stock Returns: Cross-Section and Time-Series Effects," *Journal of Financial Markets*, 5, 31–56.

Amihud, Y., and H. Mendelson, 1986, "Asset Pricing and the Bid-Ask Spread," *Journal of Financial Economics*, 17, 223–249.

Annaert, J., F. Van Holle, J. Crombez, and B. Spinel, 2002, "Value and Size Effect: Now You See It, Now You Don't," *SSRN Working Paper*.

Banz, R.W., 1981, "The Relationship between Return and Market Value of Common Stocks," *Journal of Financial Economics*, 9, 3–18.

Barry, C.B., E. Goldreyer, L. Lockwood, and M. Rodriguez, 2001, "Robustness of Size and Value Effects in Emerging Equity Markets, 1985-2000," *SSRN Working Paper*.

Basu, S., 1983, "The Relationship between Earnings' Yield, Market Value and Return for NYSE Common Stocks: Further Evidence," *Journal of Financial Economics*, 12, 129–156.

Berges, A., J.J. McConnell, and G.G. Schlarbaum, 1984, "The Turn-of-the-Year in Canada," *Journal of Finance*, 39, 185–192.

Berk, J.B., 1995, "A Critique of Size-Related Anomalies," *Review of Financial Studies*, 8, 275–286.

Berk, J.B., 1996, "An Empirical Re-examination of the Relation between Firm Size and Return," working paper, University of Washington.

Berk, J.B., 2000, "Sorting Out Sorts," *Journal of Finance*, 55, 407–427.

Black, F., 1993, "Beta and Return," *Journal of Portfolio Management*, 8–18.

Brennan, M.J., and A. Subrahmanyam, 1996, "Market Microstructure and Asset Pricing: On the Compensation for Illiquidity in Stock Returns," *Journal of Financial Economics*, 41, 441–464.

Brounen, D., A. de Jong, and C.G. Koedijk, 2004, "Corporate Finance in Europe: Confronting Theory with Practice," *Financial Management*, 33, 71–101.

Brown, P., D.B. Keim, A.W. Kleidon, and T.A. Marsh, 1983, "Stock Return Seasonalities and the Tax-Loss Selling Hypothesis: Analysis of the Arguments and Australian Evidence," *Journal of Financial Economics*, 12, 105–127.

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.

Campbell, J.Y., J. Hilscher, and J. Szilagyi, 2005, "In Search of Distress Risk," working paper, Harvard University.

Chan, L.K.C., and N.-F. Chen, 1991, "Structural and Return Characteristics of Small and Large Firms," *Journal of Finance*, 46, 1467–1484.

Chan, L.K.C., N.-F. Chen, and D.A. Hsieh, 1985, "An Exploratory Investigation of the Firm Size Effect," *Journal of Financial Economics*, 14, 451–471.

Coghlan, H., 1988, "Small Firms versus Large on the Irish Stock Exchange: An Analysis of the Performances," *Irish Business and Administrative Research*, 9, 10–20.

Cook, T.J., and M.S. Rozeff, 1984, "Size and Earnings/Price Ratio Anomalies: One Effect or Two?," *Journal of Financial and Quantitative Analysis*, 19, 449–466.

Corniolay, C. and J. Pasquier, 1991, "CAPM, Risk Premium Seasonality and the Size Anomaly: The Swiss Case," *Journal of the French Finance Association*, 12, 23–44.

Daniel, K., and S. Titman, 1997, "Evidence on the Characteristics of Cross-Sectional Variation in Stock Returns," *Journal of Finance*, 52, 1–33.

Datar, V.T., N.Y. Naik, and R. Radcliffe, 1998, "Illiquidity and Stock Returns: An Alternative Test," *Journal of Financial Markets*, 1, 203–219.

Davis, J.L., E.F. Fama, and K.R. French, 2000, "Characteristics, Covariances, and Average Returns: 1929 to 1997," *Journal of Finance*, 55, 389–406.

DeBondt, W.F.M., and R.H. Thaler, 1985, "Does the Stock Market Overreact?," *Journal of Finance*, 40, 793–805.

Dichev, I.D., 1998, "Is the Risk of Bankruptcy a Systematic Risk?," *Journal of Finance*, 53, 1131–1147.

Dimson, E., 1979, "Risk Measurement when Shares are Subject to Infrequent Trading," *Journal of Financial Economics*, 7, 197–226.

Dimson, E., and P. Marsh, 1999, “Murphy's Law and Market Anomalies,” *Journal of Portfolio Management*, 25, 53–69.

Dimson, E., P. Marsh, and M. Staunton, 2002, *Triumph of the Optimists: 101 Years of Global Investment Returns*, Princeton University Press.

Doeswijk, R.Q. 1997, “Contrarian Investment in the Dutch Stock Market,” *De Economist*, 145, 573–198.

Downs, T.W., and R.W. Ingram, 2000, “Beta, Size, Risk, and Return,” *Journal of Financial Research*, 23, 245–260.

Drew, M.E., T. Naughton, and M. Veeraraghavan, 2003, “Firm Size, Book-to-Market Equity and Security Returns: Evidence from the Shanghai Stock Exchange,” working paper, Queensland University of Technology.

Eleswarapu, V.R., and M.R. Reinganum, 1993, “The Seasonal Behavior of the Liquidity Premium in Asset Pricing,” *Journal of Financial Economics*, 34, 373–386.

Elfakhani, S., L.J. Lockwood, and T.S. Zaher, 1998, “Small Firm and Value Effects in the Canadian Stock Market,” *Journal of Financial Research*, 21, 277–292.

Fama, E.F., and K.R. French, 1992, “The Cross-Section of Expected Stock Returns,” *Journal of Finance*, 47, 427–465.

Fama, E.F., and K.R. French, 1993, “Common Risk Factors in the Returns on Stocks and Bonds,” *Journal of Financial Economics*, 33, 3–65.

Fama, E.F., and K.R. French, 1995, “Size and Book-to-Market Factors in Earnings and Returns,” *Journal of Finance*, 50, 131–155.

Fama, E.F., and K.R. French, 1996, “Multifactor Interpretations of Asset Pricing Anomalies,” *Journal of Finance*, 51, 55–84.

Fama, E.F., and K.R. French, 2004, “The CAPM: Theory and Evidence,” *Journal of Economic Perspectives*, 18, 25–46.

Fama, E.F., and K.R. French, 2005, “The Value Premium and the CAPM,” *Journal of Finance*, forthcoming.

Fama, E.F., and J.D. MacBeth, 1973, “Risk, Return, and Equilibrium: Empirical Tests,” *Journal of Political Economy*, 81, 607–636.

Ferguson, M.F., and R.L. Shockley, 2003, “Equilibrium “Anomalies”,” *Journal of Finance*, 58, 2549–2580.

Ferson, W.E., and C.R. Harvey, 1999, "Conditioning Variables and the Cross Section of Stock Returns," *Journal of Finance*, 54, 1325–1360.

Ferson, W.E., S. Sarkissian, and T. Simin, 1999, "The Alpha Factor Asset Pricing Model: A Parable," *Journal of Financial Markets*, 2, 49–68.

Gillan, F.L., 1990, "An Investigation into CAPM Anomalies in New Zealand: The Small Firm and Price-Earnings Ratio Effects," *Asia Pacific Journal of Management*, 7, 63–78.

Goyal, A., and P. Santa-Clara, 2003, "Idiosyncratic Risk Matters!," *Journal of Finance*, 58, 975–1007.

Graham, J.R., and C.R. Harvey, 2001, "The Theory and Practice of Corporate Finance: Evidence From the Field," *Journal of Financial Economics*, 60, 187–243.

Handa, P., S.P. Kothari, and C. Wasley, 1989, "The Relation between the Return Interval and Betas: Implications for the Size Effect," *Journal of Financial Economics*, 23, 79–100.

Hawawini, G., and D.B. Keim, 1995, "On the Predictability of Common Stock Returns: World-Wide Evidence," in R. Jarrow, V. Maksimovic, and W.T. Ziemba, (eds.), *Handbooks in Operations Research and Management Science*, North Holland, Amsterdam.

Hawawini, G., and D.B. Keim, 2000, "The Cross-Section of Common Stock Returns: A Review of the Evidence and Some New Findings," in D.B. Keim, and W.T Ziemba (eds.), *Security Market Imperfections in World-Wide Equity Markets*, Cambridge University Press.

Hawawini, G., P. Michel, and A. Corhay, 1989, "A Look at the Validity of the CAPM in Light of Equity Market Anomalies: The Case of Belgian Common Stocks," in S. Taylor (ed.), *A Reappraisal of the Efficiency of Financial Markets*, NATO ASI Series, Springer Verlag.

Herrera, M.J., and L.J. Lockwood, 1994, "The Size Effect in the Mexican Stock Market," *Journal of Banking and Finance*, 18, 621–632.

Heston, S.L., K.G. Rouwenhorst, and R.E. Wessels, 1999, "The Role of Beta and Size in the Cross-Section of European Stock Returns," *European Financial Management*, 5, 9–27.

Horowitz, J.L., T. Loughran, and N.E. Savin, 2000, "Three Analyses of the Firm Size Premium," *Journal of Empirical Finance*, 7, 143–153.

Ibbotson Associates and Standard & Poor's, 2003, *Corporate Value Consulting Risk Premium Report*.

Jagannathan, R., and Z. Wang, 1996, "The Conditional CAPM and the Cross-Section of Expected Returns," *Journal of Finance*, 51, 3–53.

Kandel, S., and R.F. Stambaugh, 1995, "Portfolio Inefficiency and the Cross-Section of Expected Returns," *Journal of Finance*, 50, 157–184.

Kato, K., and J.S. Schallheim, 1985, "Seasonal and Size Anomalies in the Japanese Stock Market," *Journal of Financial and Quantitative Analysis*, 20, 243–260.

Keim, D.B., 1983, "Size-Related Anomalies and Stock Return Seasonality: Further Empirical Evidence," *Journal of Financial Economics*, 12, 13–32.

Kim, Y.G., K.H. Chung, and C.S. Pyun, 1992, "Size, Price-Earnings Ratio, and Seasonal Anomalies in the Korean Stock Market," in S.G. Rhee and R.P. Chang (eds.), *Pacific-Basin Capital Markets Research*, Vol. III, Elsevier, 303–319.

Knez, P.J., and M.J. Ready, 1997, "On the Robustness of Size and Book-to-Market in Cross-Sectional Regressions," *Journal of Finance*, 52, 1355–1382.

Kothari, S.R., J. Shanken, and R.G. Sloan, 1995, "Another Look at the Cross-Section of Expected Stock Returns," *Journal of Finance*, 50, 185–224.

Lakonishok, J., A. Shleifer, and R.W. Vishny, 1992, "The Structure and Performance of the Money Management Industry," *Brookings Papers on Economic Activity*, 339–391.

Lakonishok, J., A. Shleifer, and R.W. Vishny, 1994, "Contrarian Investment, Extrapolation, and Risk," *Journal of Finance*, 49, 1541–1578.

Lamoureux, C.G., and G.C. Sanger, 1989, "Firm Size and Turn-of-the-Year Effects in the OTC/NASDAQ Market," *Journal of Finance*, 44, 1219–1245.

Lettau, M., and S. Ludvigson, 2001, "Resurrecting the (C)CAPM: A Cross-Sectional Test When Risk Premia Are Time-Varying," *Journal of Political Economy*, 109, 1238–1287.

Levis, M., 1985, "Are Small Firms Big Performers?," *The Investment Analyst*, 76, 21–27.

Lewellen, J., and S. Nagel, 2005, "The Conditional CAPM Does Not Explain Asset-Pricing Anomalies," *Journal of Financial Economics*, forthcoming.

Lo, A.W., and A.C. MacKinlay, 1990, "Data-Snooping Biases in Tests of Financial Asset Pricing Models," *Review of Financial Studies*, 3, 431–467.

Loughran, T., 1997, "Book-to-Market across Firm Size, Exchange, and Seasonality: Is There an Effect?," *Journal of Financial and Quantitative Analysis*, 32, 249–268.

Louvet, P., and O. Taramasco, 1991, “L'Effet “Jour-de-la-Semaine” à la Bourse de Paris: Un Effet Transactionnel,” *Journal de la Société Statistique de Paris*, 2, p. 50–76.

Ma, T., and T.Y. Shaw, 1990, “The Relationships between Market Value, P/E Ratio, and Trading Volume and the Stock Return of Taiwan Stock Exchange,” in S.G. Rhee and R.P. Chang (eds.), *Pacific-Basin Capital Markets Research*, Vol. I, Elsevier, 313–335.

MacKinlay, A.C., 1995, “Multifactor Models Do Not Explain Deviations from the CAPM,” *Journal of Financial Economics*, 38, 3–28.

Malkiel, B.G., and Y. Xu, 1997, “Risk and Return Revisited,” *Journal of Portfolio Management*, 9–14.

Malkiel, B.G., and Y. Xu, 2004, “Idiosyncratic Risk and Security Returns,” working paper, Princeton University.

Merton, R.C., 1973, “An Intertemporal Capital Asset Pricing Model,” *Econometrica*, 41, 867–887.

Merton, R.C., 1987, “A Simple Model of Capital Market Equilibrium with Incomplete Information,” *Journal of Finance*, 42, 483–510.

Pastor, L., and R.F. Stambaugh, 2003, “Liquidity Risk and Expected Stock Returns,” *Journal of Political Economy*, 111, 642-685.

Reinganum, M.R., 1981, “Misspecification of Asset Pricing: Empirical Anomalies Based on Earnings’ Yields and Market Values,” *Journal of Financial Economics*, 9, 19–46.

Reinganum, M.R., 1999, “The Significance of Market Capitalization in Portfolio Management over Time,” *Journal of Portfolio Management*, 25, 39–50.

Roll, R., 1977, “A Critique of the Asset Pricing Theory’s Test: Part I: On Past and Potential Testability of the Theory,” *Journal of Financial Economics*, 4, 129–176.

Roll, R., 1983, “Vas ist Das? The Turn-of-the-Year Effect and the Return Premia of Small Firms,” *Journal of Portfolio Management*, 9, 18–28.

Roll, R., and S. Ross, 1994, “On the Cross-sectional Relation between Expected Returns and Betas,” *Journal of Finance*, 49, 101–121.

Ross, S., 1976, “The Arbitrage Theory of Capital Asset Pricing,” *Journal of Economic Theory*, 13, 341–360.

Ross, S., 1977, "The Capital Asset Pricing Model (CAPM), Short-Sale Restrictions and Related Issues," *Journal of Finance*, 32, 177–183.

Rubio, G., 1988, "Further International Evidence on Asset Pricing: The Case of the Spanish Capital Market", *Journal of Banking and Finance*, 12, 221–242.

Santos T., and P. Veronesi, 2005, "Labor Income and Predictable Stock Returns," *Review of Financial Studies*, forthcoming.

Scholes, M., and J. Williams, 1977, "Estimating Betas from Nonsynchronous Data," *Journal of Financial Economics*, 5, 309–327.

Schultz, P., 1983, "Transaction Costs and the Small Firm Effect: A Comment", *Journal of Financial Economics*, 12, 81–88.

Schwert, G.W., 1983, "Size and Stock Returns, and Other Empirical Regularities," *Journal of Financial Economics*, 12, 3–12.

Shumway, T., and V.A. Warther, 1999, "The Delisting Bias in CRSP's Nasdaq Data and Its Implications for the Size Effect," *Journal of Finance*, 54, 2361–2379.

Stambaugh, R.F., 1982, "On the Exclusion of Assets from Tests of the Two-Parameter Model: A Sensitivity Analysis," *Journal of Financial Economics*, 10, 237–268.

Stehle, R., 1997, "Der Size-Effekt am Deutschen Aktienmarkt," *Zeitschrift für Bankrecht und Bankwirtschaft*, 9, 237–260.

Stoll, H.R., and R.E. Whaley, 1983, "Transaction Costs and the Small Firm Effect," *Journal of Financial Economics*, 12, 57–79.

Strong, N., and G. Xu, 1997, "Explaining the Cross-Section of UK Expected Stock Returns," *British Accounting Review*, 29, 1–23.

Vassalou, M., and Y. Xing, 2004, "Default Risk in Equity Returns," *Journal of Finance*, 59, 831–868.

Wahlroos, B., and T. Berglund, 1986, "Anomalies and Equilibrium Returns in a Small Stock Market," *Journal of Business Research*, 14, 423–440.

Wong, P.L., S.K. Neoh, and K.H. Lee, 1990, "Seasonality in the Malaysian Stock Market," *Asia Pacific Journal of Management*, 7, 43–62.

Ziemba, W.T., 1991, "Japanese Security Market Regularities: Monthly, Turn-of-the-Month and Year, Holiday and Golden Week Effects," *Japan and the World Economy*, 3, 119–146.

Table 1: Evidence on the size effect for the U.K. and the U.S.

This table presents an overview of the results of various empirical studies on the size effect in the U.K. and the U.S. The results of all studies presented in the table are based on univariate sorting procedures on the basis of the market value of individual stocks, except for Dimson and Marsh (1999), who use index data for the U.K. and the U.S. The columns present the estimated size premium (in percent per month), the sample period, the number of securities studied, the number of size portfolios into which the securities are sorted, the average market value of the firms in the largest size portfolio relative to the average market value of the firms in the smallest size portfolio, and the average monthly return and beta estimate of the firms in the largest and the smallest size portfolio. The estimates of the size premium are based on size-sorted portfolio returns that are unadjusted for risk.

Country Study	**Size Premium (% p.m.)**	**Test Period**	**# Securities**	**# Portfolios**	**MV Largest / Smallest**	**Return Smallest (% p.m.)**	**Return Largest (% p.m.)**	**Risk (β) Smallest**	**Risk (β) Largest**
United Kingdom									
Levis (1985)°	0.40	1958-1982	LSE	10	182	1.33	0.94	0.64	1.02
Strong and Xu (1995)°	0.61	1973-1992	NA	10	182	NA	NA	NA	NA
Dimson & Marsh (1999)	0.49	1955-1988	indices	–	–	–	–	–	–
Dimson & Marsh (1999)	–0.47	1988-1997	indices	–	–	–	–	–	–
United States									
Banz (1981)	0.4	1936-1975	NYSE	5	NA	NA	NA	NA	NA
Reinganum (1981)	1.6	1975-1977	566	10	212	1.0	–0.6	1.00	0.83
Brown, Kleidon, & Marsh (1983)	1.8	1967-1979	566	10	NA	1.2	–0.6	NA	NA
Keim (1983)	2.4	1963-1979	1500-2400	10	248	1.6	–0.8	1.47	0.98
Handa, Kothari, & Wasley (1989)	1.52	1941-1982	NYSE/AMEX	20	NA	2.40	0.88	1.41	0.67
Lamoureux & Sanger (1989)	1.9	1973-1985	7659 Nasdaq	20	449	1.6	–0.3	0.69	1.00
Lamoureux & Sanger (1989)	1.2	1973-1985	4170 NYSE/AMEX	20	1519	0.8	–0.4	0.95	0.91
Fama & French (1992)	0.74	1962-1989	NYSE/AMEX/Nasdaq	10	NA	1.64	0.90	1.44	0.90
Dimson & Marsh (1999)	0.34	1955-1983	indices	–	–	–	–	–	–
Dimson & Marsh (1999)	–0.20	1983-1997	indices	–	–	–	–	–	–

° information taken from Hawawini & Keim (1995, 2000)

Table 2: International evidence on the size effect

This table presents an overview of the results of various empirical studies on the size effect in international equity markets. Sources: Australia: Beedles (1992) [CRIF is the Center for Research in Finance]; Belgium: Hawawini, Michel, and Corhay (1989); Canada: Elfakhani, Lockwood, and Zaher (1998); China: Drew, Naughton, and Veeraraghavan (2003); Emerging markets: Barry, Goldreyer, Lockwood, and Rodriguez (2001); Europe: Annaert, Van Holle, Crombez, and Spinel (2002); Finland: Wahlroos and Berglund (1986); France: Louvet and Taramasco (1991); Germany: Stehle (1997); Ireland: Coghlan (1988); Japan: Chan, Hamao, and Lakonishok (1991); Korea: Kim, Chung, and Pyun (1992); Mexico: Herrera and Lockwood (1994); Netherlands: Doeswijk (1997); New Zealand: Gillan (1990); Singapore: Wong, Neoh, and Lee (1990); Spain: Rubio (1986); Switzerland: Cornioley and Pasquier (1991); Taiwan: Ma and Shaw (1990); Turkey: Aksu and Onder (2003).

Country	Size Premium (% p.m.)	Test Period	# Securities	# Portfolios	MV Largest / Smallest	Return Smallest (% p.m.)	Return Largest (% p.m.)	Risk (β) Smallest	Risk (β) Largest
Australia	5.06	1974-1987	CRIF data	10	267	6.82	0.76	1.43	1.04
Belgium°	0.52	1969-1983	170	5	188	1.17	0.65	1.01	0.98
Canada	0.98	1975-1992	694	5	178	2.00	1.02	0.58	0.60
China	0.92	1993-2000	44-701	2	NA	NA	NA	NA	NA
Emerging markets	2.20	1985-2000	2000	5	NA	2.94	0.74	NA	NA
Europe	1.45	1974-2000	2866	10	196	2.64	1.19	1.06	0.94
Finland	0.76	1970-1981	50	10	133	1.65	0.89	0.52	0.95
France°	0.90	1977-1988	529-460	5	NA	1.20	0.30	NA	NA
Germany	0.49	1954-1990	All FSE	9	NA	1.54	1.05	0.80	1.08
Ireland	0.47	1977-1986	40	5	NA	3.10	2.63	NA	NA
Japan	0.97	1971-1988	1570	4	57	2.44	1.47	1.10	0.81
Korea°	–0.40	1984-1988	NA	10	62	NA	NA	NA	NA
Mexico	4.16	1987-1992	100	3	37	5.80	1.64	1.31	0.79
Netherlands	0.13#	1973-1995	145-183	5	NA	NA	NA	NA	NA
New Zealand°	0.51§	1977-1984	100	5	60	0.69	0.18	0.90	0.99
Singapore°	0.41	1975-1985	NA	3	23	NA	NA	NA	NA
Spain	0.56§	1963-1982	160	10	228	0.58	0.02	NA	NA
Switzerland°	0.52	1973-1988	153	6	99	0.94	0.42	NA	NA
Taiwan°	0.57	1979-1986	53-72	5	17	0.47	–0.10	0.55	0.72
Turkey	3.42	1993-1997	106-173	4	NA	NA	NA	NA	NA

° information taken from Hawawini & Keim (1995, 2000); # not significant; § based on risk-adjusted returns (see footnote 6 of this paper)

Figure 1: The size effect in U.S. equity returns 1927-2004

This figure depicts the annual, market-weighted return differential between the smallest and the largest size quintile of all NYSE, AMEX, and Nasdaq stocks over the period 1927-2004. These returns are not adjusted for (market) risk.

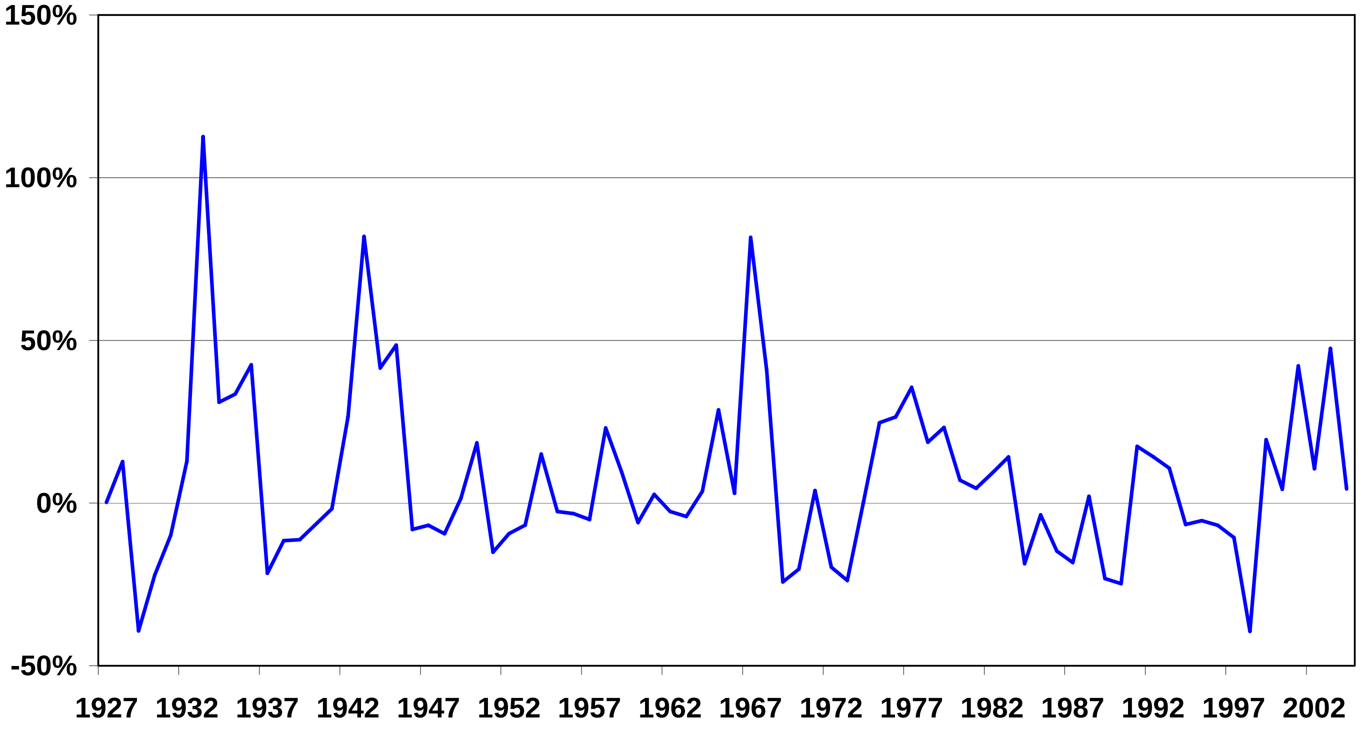

Figure 2: The returns on the market, SMB, and HML portfolio 1927-2004

This figure shows the annual returns on the U.S. market portfolio (in excess of the risk-free rate) and the SMB and HML portfolios over the period 1927-2004.

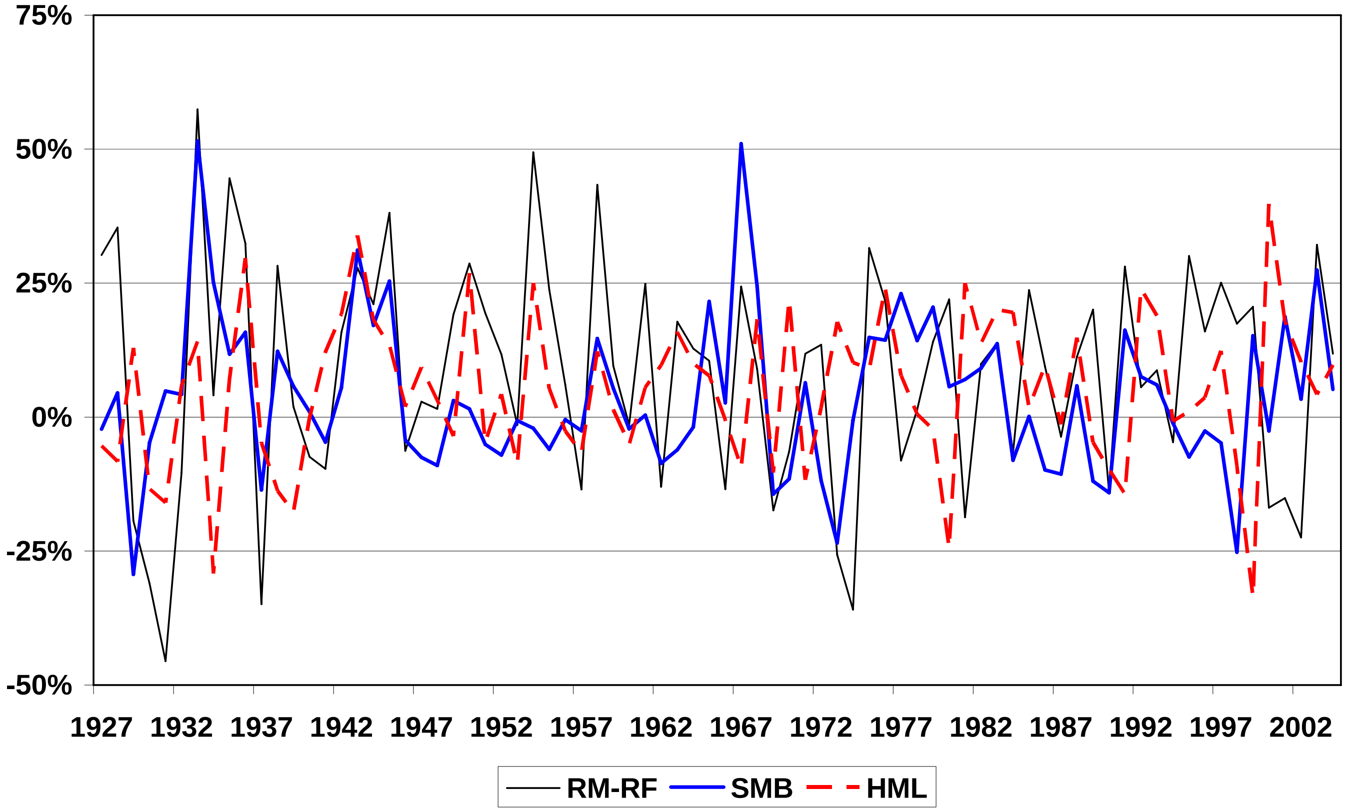

Figure 3: Seasonal patterns in the size effect in U.S. equity returns 1927-2004

This figure depicts the average market-weighted return differential between the smallest and the largest size-quintile of all NYSE, AMEX, and Nasdaq stocks in each month over the period 1927-2004. These returns are not adjusted for (market) risk.