Firm Size and Stock Returns:
A Meta-Analysis

Anton Astakhov, Tomas Havranek, and Jiri Novak

Czech National Bank and Charles University, Prague

Abstract

A prominent factor used in most models predicting stock returns is firm size. Yet no consensus has emerged on the magnitude and stability of the size premium, with some researchers even questioning the usefulness of the factor. To take stock of the voluminous academic literature on the size premium, we collect 1,746 estimates of the effect of size on returns reported in 102 published studies and conduct the first meta-analysis on this topic. We find evidence of strong publication bias: researchers prefer to report estimates that are statistically significant and show a negative relation between size and returns, exaggerating the mean reported coefficient threefold. After correcting for the bias, we find that the literature suggests a size premium (the difference in annual stock returns on the smallest and largest capitalization quintile) of 1.72%. We also find that the premium was much larger prior to the publication of the first study on the topic. Moreover, we show that the intensity of publication bias has been decreasing over time.

Keywords

asset pricing; stock returns; risk; size premium; multi-factor models; publication selection bias

JEL Classification

*C*18, *G*12, *G*15

The data and code used in this study are available in an online appendix at http://meta-analysis.cz/size. Corresponding author: Anton Astakhov, av.astakhov@gmail.com.

1. Introduction

The identification of factors explaining the variation in stock returns and estimating the magnitude of the corresponding return premiums is a fundamental research agenda in asset pricing. It is commonly assumed that on the margin, investors are risk averse and set asset prices so that the implied expected returns provide a fair compensation for the systematic risk that each asset involves. The Capital Asset Pricing Model (CAPM; Sharpe 1964; Lintner 1965; Black 1972) presumes that an asset's systematic risk depends on its contribution to the volatility of returns on the market portfolio. The model approximates an asset's systematic risk by its market beta, i.e., by the sensitivity of its returns to the variation in market returns, and predicts a positive linear association between expected return on an asset and its market beta. The model also suggests that the market beta is a compressive measure of an asset's systematic risk, and so no other factor should be relevant for explaining variation in stock returns.

Contrary to this prediction, empirical research identifies numerous firm characteristics associated with realized stock returns, e.g., firm size (Banz 1981; Reinganum 1981; Keim 1983), financial leverage (Bhandari 1988), earnings-to-price ratio (Reinganum 1981; Basu 1983), book-to-market equity ratio (B/M; Fama and French 1992, 1996; Lakonishok et al. 1994), and stock price momentum (Jegadeesh and Titman 1993, 2001; Carhart 1997). These findings suggest that systematic risk may be a more elusive concept than envisaged by the CAPM. Risk can plausibly be multi-dimensional and the above variables may predict stock returns because they capture a firm's exposure to various underlying risk dimensions that are disregarded in the CAPM framework, such as bankruptcy, liquidity, and information risks.

We perform a meta-analysis of the published studies on the association between firm size and realized stock returns. Firm size is one of the first empirically documented characteristics associated with realized stock returns (Banz 1981; Reinganum 1981; Keim

1983). Chen et al. (1986) argue that *"size may be the best theory we now have of expected returns"* (p. 394). The "size effect", i.e., the tendency of low market capitalization stocks (i.e., "small stocks") to yield higher returns than high market capitalization stocks (i.e., "large stocks"), has been subsequently tested in a wide range of markets and over different time periods. Firm size is used as a factor in all major empirical pricing models, i.e., in the three-factor model (Fama and French 1993), in the four-factor model (Carhart 1997), and in the five-factor model (Fama and French 2015). These models constitute a *de facto* industry standard for estimating expected return on equity. Furthermore, survey evidence suggests that firm size is one of the most common characteristics that chief financial officers (CFOs) consider when estimating the discount rate for investment appraisal (Graham and Harvey 2001, Table 4). Thus, the effect has a profound impact on the cost of capital estimation in financial practice.

Despite its widespread recognition as a pricing factor, firm size exhibits several puzzling features that are difficult to reconcile with its ability to proxy for underlying risk. The size premium, i.e., the difference between average stock returns earned on small and large stocks, does not seem consistent across regions (Hou et al. 2011; Cakici et al. 2016) and across time (van Dijk 2011; Joel L. Horowitz et al. 2000). It is concentrated in the month of January (Keim 1983) and in the very smallest stocks (Banz 1981); when removing 1 percent of the observations of the smallest stocks, the negative association between firm size and realized returns often reverses (Knez and Ready 1997). Contrary to the notion that risky assets yield low returns in times of economic downturn, the size premium is concentrated in periods of general stock market decline (Hur et al. 2014). Furthermore, it is not obvious why large stocks should be considered less risky while stocks that have recently increased their market capitalization, i.e., stocks with a positive stock price momentum, are seen as more risky (Conrad and Kaul 1998; Chordia and Shivakumar 2002; Carhart 1997).

These puzzling findings give rise to the controversy over whether firm size indeed proxies for a firm's exposure to underlying risk or whether the findings on its predictive power reflect (i) its correlation with costly imperfections in the stock market microstructure, (ii) its ability to proxy for systematic mispricing that is subsequently corrected, or (iii) incorrect statistical inferences based on non-representative samples, inappropriate methodology, or other sources of bias (van Dijk 2011). Discriminating among these alternative explanations has profound implications for the interpretation of the findings. Thus, to understand the underlying reasons for the documented association and to judge which risk dimension (if any) it captures, it is important to study them systematically and to assess the consistency of their predictive power across different regions and over various time periods.

Unsurprisingly, prior research has amassed an extensive body of empirical evidence about the relevance of the firm size in various settings and time periods (for an overview, see Harvey et al. 2016). These findings are often contradictory and context dependent, which complicates their generalization and adoption in financial practice. Hence, it is important to synthesize the existing evidence on the size effect to provide insights that are helpful for explaining the underlying reasons for the ability of firm size to predict stock returns. Lee (2001) suggests: *"I believe future studies along these lines will not merely document new anomalies, but will also help to explain them"* (p. 242).

In line with the above suggestion, performing a meta-analysis allows us to systematically aggregate and synthesize the entire pool of past findings, to analyze the consistency of the reported estimates across various contexts, to evaluate the relative plausibility of the alternative explanations, and to assess the validity of firm size as a pricing factor, which is important for a general understanding of why firm size affects asset prices and that has important implications for the development of finance theory. Furthermore, the meta-analysis approach allows us to adjust for potential publication selection bias among the reported estimates and to estimate the magnitude of the size premium in a way that avoids the criticized pre-sorting of stocks into portfolios based on previously empirically documented characteristics, which is important for practical finance application, including the estimation of the cost of capital and the performance evaluation of mutual funds.

Indeed, we find that the existing literature suffers from significant publication selection bias, which implies that the conventional estimates of the size risk premium are significantly overstated. Our estimate of the difference in stock returns on the smallest and largest market capitalization quintile is 1.72% after controlling for the effects of publication selection. This contrasts with the mean reported size premium of approximately 5%. Surprisingly, the quality of the journal does not seem to affect the magnitude of the publication bias. Nevertheless, the magnitude of the bias is associated with the publication year of the study, with newer studies being less involved in publication selection.

The decrease in publication bias might be explained by the fact that more and more researchers have become convinced that the size premium has moderated after the 1970s (particularly after the publication of the first study on the topic, Banz 1981). We show a drop of approximately 50% in estimated size premiums and thus support the findings of McLean and Pontiff (2016) who show that a premium associated with measures that explain cross-sectional variation in stock returns decreases after the predictive power of the factor is first reported in empirical research and so subsequent studies tend to document lower estimates for the factor's predictive power. The growing acceptance of a decreased size premium makes it easier for newer studies to report nonsignificant (or even positive) estimates of the size effect. Although we document regional variation in the reported size premiums, the differences are not large. In contrast, we support the notion that the size premium is concentrated in the month of January.

The remainder of the paper is organized as follows. Section 2 reviews some of the related literature, Section 3 discusses methodology and the data sample, Section 4 presents the empirical results, and Section 5 concludes the paper. Appendix A lists the studies included in the meta-analysis; Appendix B presents additional results to our examination of the mediating factors of publication bias.

2. Related Literature

Even though firm size is widely used as a predictor of future stock returns, the interpretation of the findings on the size premium continues to pose a challenge in asset pricing. The use of firm size as a pricing factor and its inclusion in asset pricing models hinges on the assumption that firm size proxies for an underlying risk characteristic. Nevertheless, the relationship of firm size to systematic risk had not been *ex ante* anticipated by financial theory; instead, the conceptual underpinnings were provided only *ex post* after the association with the ability of firm size to predict stock returns was documented empirically. Thus, there is a controversy over whether the empirical findings on the firm size premium should be interpreted as evidence of the ability to proxy for underlying risk characteristics.

The size premium exhibits several puzzling characteristics that fuel this skepticism, as briefly mentioned in the previous section. First, the size premium does not seem to be constant across regions. Hou et al. (2011) investigate stock returns in 49 economies, reject the covariance risk model of firm size, and conclude that the stock price momentum and the ratio of cash flow-to-price best capture the variation in stock returns. Cakici, Tang, and Yan (2016) examine the pricing factors in 18 emerging stock markets and conclude that both firm size and momentum fail to reliably predict future stock returns. These findings are problematic because if firm size proxies for an underlying risk dimension, we would expect it to be systematically associated with stock returns in various contexts. Second, some studies suggest that the size premium disappeared after the 1980s and then resurfaced after 2000 (van Dijk 2011). Proxies for fundamental risk dimensions should be relatively persistent over time. Horowitz et al. (2000) view the disappearing size effect as evidence that firm size should not be interpreted as a proxy for a hidden risk factor.

Third, the premium is concentrated in the month of January (Keim 1983). It is not trivial to explain why small stocks are systematically riskier in January, which would give rise to the size premium compensating for the higher risk while they are no risker than other stocks in the remaining eleven months. Fourth, the size premium is positive when there is a general tendency of stock prices to decrease (i.e., in the "down markets" or the "bear markets"), whereas it is close to zero when the stock prices have a tendency to rise (i.e., in the "up markets" or the "bull markets"; Hur et al. 2014). This seems to contrast the notion of systematic risk, which entails low returns on riskier assets in times of economic downturn when scarcity of income is high.

Fifth, the premium is concentrated in the smallest stocks (Banz 1981). Knez and Ready (1997) argue that the size effect is driven by the extreme 1 percent of the observations; when these are eliminated, the negative association between firm size and realized returns reverses. The non-linearity begs an explanation of why the smallest stocks are riskier than the remaining stocks while at the same time within the remaining stocks, larger stocks are riskier than medium-sized stocks. Sixth, if driven by underlying risk the size premium is difficult to reconcile for the premium for stocks with positive "stock price momentum" (i.e., the past 6-month dividend adjusted stock return; Conrad and Kaul 1998; Chordia and Shivakumar 2002; Carhart 1997). A positive stock price momentum implies an increase in a firm's market capitalization, i.e., in firm size. It is challenging to reconcile why small stocks are riskier while in contrast, stocks that have recently been increasing in size (i.e., stocks with a positive stock price momentum) are less risky. More generally, researchers have noted the shortcomings of the entire modeling framework. As MacKinlay (1995) explains:

> *"A number of studies have presented evidence rejecting the validity of the Sharpe-Lintner capital asset pricing model (CAPM). Possible alternatives include risk-based models, such as multifactor asset pricing models, or non-risk-based models which address biases in empirical methodology, the existence of market frictions, or the presence of irrational investors. Distinguishing between the alternatives is important for applications such as cost of capital estimation"* (Abstract).

First, the CAPM is based on a number of simplifying assumptions that are needed to make the problem tractable. These assumptions are violated in real life and so the model suffers from "assuming away" important sources of risk. For example, the CAPM assumes that all information is readily available and costless to process. Realistically, collecting information and processing it for the sake of a business decision may involve significant costs. Since a part of the information production cost tends to be fixed, small firms are likely to provide less informative disclosures than large firms. Because the risk resulting from a firm's opacity may not be fully diversifiable, higher realized returns on small stocks reflect compensation for the higher information risk (Barry and Brown 1984). In a similar vein, the CAPM assumes that financial distress and bankruptcy are costless even though they actually entail significant costs. If a company is marginalized because it is unable to keep up with the industry level of production efficiency, its stock price is likely to be hammered, depressing the stock's market capitalization. Higher realized stock returns on small stocks may thus reflect compensation for the higher risk of financial distress (Chan and Chen 1991; Vassalou and Xing 2004).

Second, the CAPM assumes a frictionless market in which trading is costless and individual trades have no price impact. In reality, assets vary in terms of their trading costs and market depth. Small stocks have higher transaction costs and the market for them may be shallower, because fewer investors tend to be interested in being a counterparty in these transactions. Thus, small stocks may be less liquid, and some investors (e.g., institutional investors) may prefer to stay away from them (Amihud 2002). Hence, the size premium may reflect a compensation for market imperfections that impair the liquidity of small stocks.

Third, the CAPM assumes that the financial market is efficient, investors rationally evaluate all available information, and assets are fairly priced in that the established price is an unbiased estimate of a firm's intrinsic value based on the information available at that point in time. Behavioral finance research suggests that investors are not always perfectly rational, and temporary departures from efficient pricing may occur (e.g., Lakonishok et al. 1994). If firm size is correlated with systematic mispricing, the excess returns may reflect stock price corrections rather than compensate for higher risk. For example, small stocks may see their prices temporarily depressed because the stock market has overreacted to a stream of bad news. In such a case, the prices of small stocks would be expected to increase in the future as the market gradually corrects the mispricing.

Fourth, it is possible that the reported associations are spurious, a result of flawed methodology, data mining, extreme observations, or other sources of bias. For example, Lo and MacKinlay (1990), MacKinlay (1995), and Berk (2000) argue when stocks are sorted into portfolios based on previously empirically documented characteristics, conventional tests may overstate statistical significance. Furthermore, Knez and Ready (1997) argue that the size effect is driven by the extreme 1 percent of the observations. When 1 percent of the most extreme observations is trimmed, the size effect is reversed.

Distinguishing between these alternative explanations for the predictive ability of firm size is crucial both for a general understanding of how financial markets set prices and for practical finance application, including the estimation of the cost of capital and the performance evaluation of mutual funds. If the superior return on small stocks represents a compensation for higher distress risk, then firm size (or an asset's sensitivity to the small-minus-big (SMB) factor; Fama and French 1992, 1993) is a relevant determinant of a firm's cost of capital. If instead small stocks generate higher returns because they are less liquid, then firm size is relevant for determining the required return of large institutional investors, but not necessarily for small individual investors who trade smaller volumes and therefore are less likely to be adversely affected by impaired stock liquidity. In contrast, if superior returns on small stocks reflect a correction of temporary mispricing or if the reported findings are based on an inappropriate methodology, then firm size should be disregarded in computing the cost of capital.

3. Research Design

3.1. Methodology

The size effect has traditionally been estimated using two main approaches (Crain 2011; Fama and French 2008). First, Banz (1981), along with numerous subsequent studies, employs a cross-sectional regression of (excess) individual or portfolio-sorted stock returns on a measure of size. A typical proxy for size is the market value of equity (the number of shares outstanding times share price); since this measure is bounded by zero and right-skewed, researchers tend to log-transform it. Another regression specification often used in the literature follows Fama and French (1992), in which returns are regressed on factors rather than firm-specific variables, such as in the Fama-French three-factor model. In this case, the size effect is measured by the sign and magnitude of the regression slope for the variable SMB, which represents excess returns of a small market cap stocks portfolio over a large market cap stocks portfolio.

The second way to estimate the size effect, employed, e.g., in Easterday et al. (2009), Horowitz et al. (2000), and others, is to pre-sort stocks into portfolios based on individual firm characteristics, such as size, B/M, beta, momentum, and liquidity, and directly compare the returns of the smallest and the largest quintiles of stocks. Fama and French (2008) provide a discussion of the pros and cons of portfolio sorting and regression methods.

In this paper, we consider studies that estimate the size effect by regressing excess returns on the logarithm of market value of equity or similar variables.[1] We have several reasons for doing so. First, the method featuring a direct comparison of the returns of the smallest and largest quintiles of market capitalization is unsuitable for the meta-analysis technique employed in this study, since standard errors (or t-statistics) for the difference in returns are typically not reported by the author. Fama and French (2008) provide an additional argument against using simple sorts for making inferences about the determinants of cross-sectional returns:

> *… sorts are awkward for drawing inferences about which anomaly variables have unique information about average returns. Multiple regression slopes provide direct estimates of marginal effects. Moreover, with our large samples, marginal effects are measured precisely for many explanatory variables. Second, sorts are clumsy for examining the functional form of the relation between average returns and an anomaly variable. In contrast, simple diagnostics on the regression residuals allow us to judge whether the relations between anomaly variables and average returns implied by the regression slopes show up across the full ranges of the variables.* (p.1654)

Most of the studies that estimate the size effect by regressing stock returns on market

[1] Some studies use market value of equity relative to cross-sectional average (e.g., within specific industry or geographical region). Our main results, reported in Table 3, are robust to the exclusion of these observations.

value of equity employ the so-called two-pass regression technique, developed by Fama and MacBeth (1973) (henceforth, the FM regression). Below, we include a brief description of this technique, drawing heavily on Fletcher (1997). In the first step, market β is estimated from equation (2).

$$R_{it} = \alpha_i + \beta_i R_{mt} + \epsilon_{it}, \tag{2}$$

where R_{it} – return on stock i in period t,
R_{mt} – return on market portfolio in period t,
β_i – estimated beta of stock i,
ϵ_{it} – random error term.

The second stage envisages the estimation of monthly cross-section regression (3), which can also be expanded with additional variables, which are believed to be correlated with (excess) returns, such as size:

$$R_{it} = \gamma_{0t} + \gamma_{1t}\beta_i + \gamma_{2t}X_i + u_{it}, \tag{3}$$

where β_i – beta estimated from (2),
X_i – vector of additional explanatory variables (including size).

In this or a similar form, the two-pass technique has been employed in a vast number of studies addressing the cross-sectional determinants of stock returns in general (a recent overview of such studies is available at Harvey et al. 2016) and with the size effect in particular. We do not, however, restrict our sample to those studies that employ this technique, but instead include papers reporting any form of regression of returns on size.

Having collected point estimates and standard errors from the studies that satisfy the criteria described above, we test for the selective reporting bias. Many authors have recognized this bias as a serious issue in economics research (Ashenfelter and Greenstone 2004; Stanley

2005; Havranek 2015) that arises from researchers' tendency to primarily report the estimates that display the sign predicted by the theory, along with those that are statistically significant. More formally, in the presence of selective reporting bias, the reported estimates will be correlated with their standard errors, and coefficient σ in regression (4) will be negative and significant.

If there is no publication bias, σ should be zero because the ratio of the reported estimates of the size effect to their standard errors has a t-distribution (this property is guaranteed by the econometric techniques used by the researchers estimating the size effect). If, in contrast, statistically significant estimates are preferentially reported, researchers need large point estimates to offset standard errors, giving rise to a non-zero σ; similarly, a preference for a particular sign of the estimates will lead to a non-zero σ because of the heteroskedasticity of the regression. A test of the hypothesis $H_0\colon \sigma = 0$ is known as the funnel-asymmetry test because it can also be examined visually by evaluating the corresponding scatter plot, in which symmetry is the key feature (Stanley 2005; Feld and Heckemeyer 2011; Feld et al. 2013). Coefficient S_0 from regression (4) represents the underlying mean effect corrected for the reporting bias:

$$S_{it} = S_0 + \sigma \cdot SE(S_{it}) + \epsilon_{it}, \tag{4}$$

where S_{it} – i-th estimate of size effect (that is, coefficient $\hat{\gamma}_{2t}$ from regression (3)) reported in study j.

We estimate (4) using several specifications. First, we use OLS with standard errors clustered at the level of individual studies and geographical regions. Second, we run a panel data regression using both study-fixed effects and between effects. In addition, we follow Stanley (2005, 2008) by estimating regression (4) using weighted least squares (WLS) to give more weight to more precise studies and to directly address heteroskedasticity. In this specification, the regression takes the form shown in (5), with variable $1/SE(S_{it})$, which we call *precision*, used as a weight. In addition, following Havranek et al. (2015), we run specification (5) using the inverse of the number of size effect estimates reported per study as a weight (some studies report many more estimates than other studies). The weighted regression becomes:

$$\frac{S_{it}}{SE(S_{it})} = S_0 \frac{1}{SE(S_{it})} + \sigma + u_{it} \tag{5}$$

In (5), the dependent variable is essentially the reported t-statistic, and σ measures the selective reporting bias. In the extreme case of a very strong selective reporting bias, the σ coefficient will approach -2; this would mean that the true size effect is zero, but because of the selective reporting bias, only negative and statistically significant estimates (at the 5% level) are reported in the literature.

Finally, because there is an endogeneity issue in specification (5)—some omitted study characteristics, such as estimation techniques, can easily affect both estimates and their standard errors—we also estimate an instrumental variable regression using the reciprocal of the square root of the number of observations, $1/\sqrt{n}$, as an instrument (following Zigraiova and Havranek 2016 and Havranek and Irsova 2017). This variable is correlated with the standard error, but not with the choice of the estimation technique, thus intuitively constituting a valid instrument.

Next, we also include interaction terms of the standard error with the recursive impact factor (as reported on the IDEAS/RePEc website[2]) and with the year of publication in

[2] Available at https://ideas.repec.org/top/top.journals.recurse.html. Data on the recursive impact factor were downloaded on July 8, 2017; six studies from the sample were published in outlets that did not have a recursive impact factor listed on the IDEAS/RePEc website. For an application of this measure in meta-analysis, see, for example, Havranek and Irsova (2011).

specification (4). The effect of a study's publication in a higher-impact journal on the strength of selective reporting bias is uncertain: on the one hand, higher-quality journals employ more stringent review procedures, decreasing the possibility of publishing studies with heavy data mining and backfitting of results; on the other hand, researchers may be reluctant to submit studies with inconclusive or "unexpected" results (such as nonsignificant or positive size effects) to high-impact journals because of reputational concerns and fear of rejection. The same uncertainty regarding the effect on the selective reporting bias applies to the year of publication: on the one hand, contemporary studies employ more refined econometric techniques, building on past advances, which should lead to the discovered effect being closer to the true effect; on the other hand, more recent research might well stick to the convention of the negative size effect and avoid publishing results that contradict this convention.

As an additional test of selective reporting, we employ Hedges' (1992) model for detecting publication selection bias. It models the selection process using a weight function, which is in turn approximated by a step function with discontinuities resembling the psychological "barriers" associated with interpretation of different p-values. Hedges (1992) refers to the findings of psychological studies on the interpretation of research results showing that conclusiveness of the results is strongly associated with the p-value. Moreover, this perception is distorted in such a way that perceived conclusiveness changes drastically near the conventionally used p-values of 0.1, 0.05 and 0.01; that is, the result with a p-value of, e.g., 0.045, is perceived as much more conclusive than the result with a p-value of 0.055.

Hedges (1992) uses these findings to develop a model in which the probability of selection (publication of a study) is a function of p-value, and the abovementioned discontinuities are incorporated into the step function; that is, the probability of selection changes when p-value approaches a psychological barrier (e.g., 0.1, 0.05 or 0.01).

3.2. Data Sample

We collected the dataset for meta-analysis using several approaches. First, we performed a search in Google Scholar using query *(size OR small) AROUND(4) (effect OR premium OR anomaly OR pattern OR puzzle)" size small stock firm returns risk empirical regression portfolio sort effect premium "market value" OR "market capitalization*, which returned many hits; from these, the first 500 results sorted by relevance were examined. The search was restricted to studies published in peer-reviewed journals during or after 1981, the year of Banz's (1981) original research on size effect. Because the peer-review process preceding the publication of a paper in a journal serves as a quality control mechanism, we only consider estimates reported in published studies and therefore we base our analysis on estimates that are likely to make a significant impact on the research and business community. In addition, Rusnak et al. (2013) find no difference in the extent of selective reporting between published and unpublished studies in economics; authors want to eventually publish their working papers, so the same incentives also apply to non-published research. The search for studies was terminated on March 31, 2017; the data and code are available at meta-analysis.cz/size.

We proceed by checking all the references from the ten most-cited papers in the dataset (this technique is known as "snowballing" in systematic review and meta-analysis studies), collected via Google Scholar search, along with several relevant review papers, specifically, (Crain 2011; van Dijk 2011; Harvey et al. 2016; Horowitz et al. 2000; Hou et al. 2011) and others. To be included in the dataset, a study must provide a point estimate and t-statistic or a standard error of the regression of returns on a natural logarithm of the market value of equity. We do not restrict the dataset to the studies that employ the two-pass FM regression as in (2) and (3), even though this method is essentially standard practice. Predictor and response variables are also allowed to vary in their design: returns can be simple, excess (over, e.g., the risk-free rate) or otherwise adjusted, whereas the size variable can be relative to the cross-sectional or time-series average. We collect all estimates given in individual studies. The resulting dataset contains 1,746 estimates from 102 papers, which puts this study among the largest meta-analyses conducted in economics and finance (for statistics, see Doucouliagos and Stanley 2013). A list of studies selected for meta-analysis is provided in Appendix A.

Table 1: Summary statistics

	Mean	St. dev.	Min	Max
Size coefficient	-0.092	0.482	-5.94	4.69
t-statistic	-1.246	2.956	-16.84	26.18
SE	0.109	0.339	0.00	4.86
$1/\sqrt{n}$	0.077	0.039	0.03	0.30
Start Year	1965.783	95.542	0.00	2002.00
End Year	1989.253	96.053	0.00	2011.00
# of observations	265.704	174.833	11.00	905.00
Publication Year	2000.549	7.479	1981.00	2014.00
N	1746			

The summary statistics resulting from our dataset are presented in Table 1. The mean reported estimate of coefficient $\hat{\gamma}_{2t}$ from regression (3) is -0.092, which corresponds to the intuitive notion of a negative size effect: larger companies display lower returns. Later in the analysis, we show how this translates into an actual size premium that can be used in practice.

The dataset exhibits significant heterogeneity in terms of the methodology employed, geographical region and time period covered. The mean estimate of the size effect differs depending on the period and country. As shown in Table 2, the size effect is less prominent after 1981, the year in which the first study on this topic was published, which is consistent with the evidence from van Dijk (2011) and Horowitz et al. (2000). In addition, the size effect is somewhat less pronounced in the USA than in other regions used in the data sample, although the difference is not great. A drastic difference in the magnitude and significance of the size effect can be observed between estimates made in January only and in any month other than January, in accordance with findings of Keim (1983). Finally, estimation techniques that are more advanced than OLS tend to yield a size effect of a smaller magnitude.

Table 2: Size effects for different subsets of data

	N	Mean	St. dev.	Min	Max
Year of observations					
Before 1981	134	-0.230	0.489	-2.24	0.03
1981-2000	276	-0.099	0.322	-1.84	0.92
After 1981	352	-0.124	0.358	-1.84	0.92
After 2000	16	-0.079	0.104	-0.27	0.00
Geographical region					
North America	1109	-0.079	0.475	-3.60	4.45
Europe	165	-0.139	0.685	-3.24	4.69
Other	458	-0.106	0.411	-5.94	0.92
Stock exchange (US)					
NYSE only	305	-0.031	0.123	-1.19	0.21
Any	784	-0.100	0.558	-3.60	4.45
Month of returns					
January only	113	-0.481	0.851	-5.94	0.92
Non-January	1633	-0.043	0.236	-1.80	0.52
Estimation technique					
OLS	928	-0.114	0.543	-5.94	4.69
Other	818	-0.068	0.400	-3.60	1.48
Stock returns					
Individual stock returns	1072	-0.109	0.524	-3.60	4.69
Returns on pre-sorted stock portfolios	628	-0.084	0.388	-5.94	3.20
All Estimates	1746	-0.09	0.48	-5.9	5

To alleviate the influence of extreme observations of the size effect, the dataset is winsorized at the 2.5% level, an approach that has become relatively common in meta-analysis (see, for example, Havranek et al., 2017). Apart from the size effect estimate, winsorization is also performed with respect to variables such as precision and the reciprocal of the square root of the number of observations.

It should be noted that during the data collection, we have encountered several difficulties caused by an ambiguous or incomplete description of the methodology in the studies. These include four issues:

1. *Different measures of size and return*. In some cases, size was measured relative to local, industry or time-series average. Returns used in the dataset can be simple, in excess of risk-free rate or adjusted using some sort of asset-pricing

models; in addition, the time frame over which returns were measured varies greatly from study to study, including such cases as 1-year ahead monthly returns (Dichev 1998) or 48-month buy-and-hold returns (Dissanaike 2002).

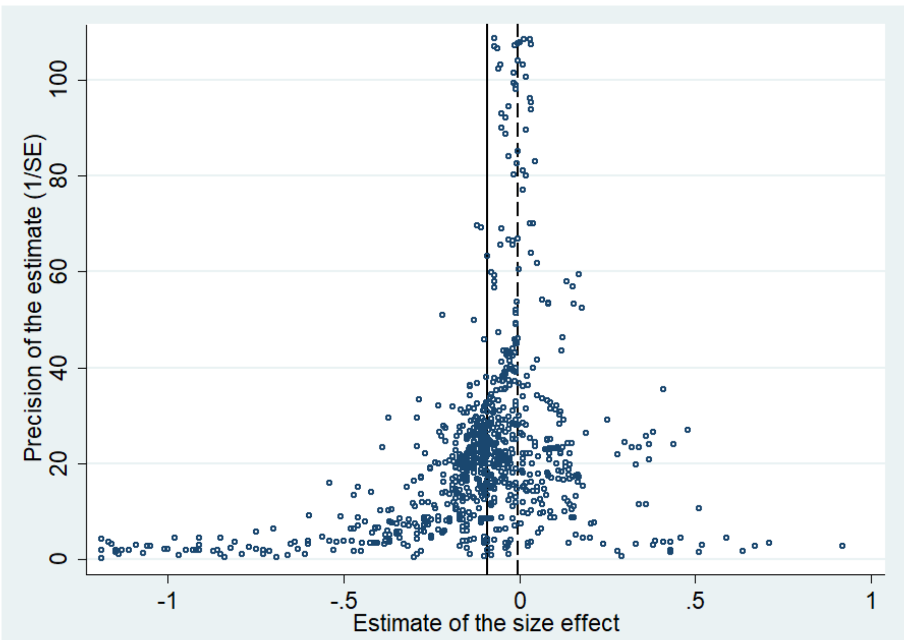
2. *Different measures of variability*. Different authors use different adjustments of standard errors and t-statistic to account for autocorrelation (Asparouhova et al. 2013), measurement errors (Fletcher 1997), and heteroscedasticity (Chen et al. 2002), *etc*.
3. *Insufficient reporting*. During the data collection, we encountered numerous cases in which certain dataset adjustments and regression outputs were not reported in a sufficiently clear manner (e.g., in some cases number of firm-level observations, scale of size variable, precise definitions of risk-free rate or portfolio sorting techniques could not be retrieved). We did our best to obtain the information required from the authors; we generally avoided dropping studies from the sample unless critical information was missing.
4. *Errors and omissions in reporting*. In certain cases, we encountered mistakes such as negative standard errors reported, standard errors apparently reported as t-statistics (and vice versa), non-matching signs of coefficient estimate and t-statistics. In addition, because of a (usually) small scale of size effect coefficient, in a few cases standard errors were reported as zeros because of rounding. Here, we assumed that the standard error was the largest possible to be still rounded at zero.

Despite the issues described above, the initial results in Tables 1 and 2 indicate not only strong evidence of the negative size effect as reported in the literature, but also its clear heterogeneity. We further investigate the issue of the underlying size effect in the following section, especially in relation to publication selection bias.

4. Results

We start our analysis with a visual investigation of potential selective reporting using the funnel plot, as suggested by Egger et al. (1997). The funnel plot depicts the precision of the estimates of the size coefficient on the *y*-axis and the point estimates of the size coefficient on the *x*-axis. In the absence of selective reporting bias, the graph should take a symmetrical funnel shape, with the most precise estimates concentrated around the underlying mean value of the size effect, whereas less precise estimates are dispersed around the mean. The funnel plot is shown in Figure 1 and presents clear evidence of selective reporting: the plot is asymmetric with point estimates concentrated in the left-hand tail. Therefore, other things being equal, positive estimates of the relation between size and returns are less likely to be reported in academic studies.

Figure 1: The funnel plot



Notes: The solid vertical line displays the sample mean; the dashed vertical line displays the sample median. Because of the presence of extreme observations for both size and precision, both variables are trimmed for better visibility.

We proceed by testing for selective reporting bias more formally using specifications (4) and (5) described in Section 3. Column (1) in Table 3 presents the baseline result of regressing the size coefficient estimate on its standard error using OLS. A negative and significant σ coefficient indicates a strong selective reporting bias, as explained in Section 3. The estimated constant of -0.0315 represents the underlying mean size effect corrected for the selective reporting bias, which is roughly 3 times smaller in absolute value than the mean of -0.092 reported in Table 1. The baseline result therefore suggests that evidence for the size effect is present in the data, but less so than the published results would suggest.

Table 3: Estimating the magnitude of the selective reporting bias

	(1) OLS	(2) FE	(3) BE	(4) Precision	(5) Study	(6) IV
SE	-0.808***	-0.897***	-0.526***	-1.159***	-0.554***	-0.494
	(0.0862)	(0.127)	(0.102)	(0.0300)	(0.165)	(0.311)
Constant	-0.0315***	-0.0236**	-0.0387	-0.000291	-0.0358***	-0.0611**
	(0.00814)	(0.0113)	(0.0245)	(0.000271)	(0.0128)	(0.0277)
N	1746	1746	1746	1746	1746	1742

Notes: The table shows the results of regression $S_{it} = S_0 + \sigma \cdot SE(S_{it}) + \epsilon_{it}$, where S_{it} – i-th estimate of size effect reported in study j and $SE(S_{it})$ is the standard error. Specification (1) is estimated using OLS with standard errors clustered by study and geographic region. Specifications (2) and (3) are panel data regressions with fixed and between effects, respectively. Specifications (4) and (5) are estimated using WLS with precision and reciprocal of number of size effect estimates reported per study as a weight. Specification (6) is the instrumental variables regression with the reciprocal of the square root of number of observations used as an instrument, with standard errors clustered by study and geographic region. Specification (6) is a panel data instrumental variables regression with fixed effects and the reciprocal of the square root of the number of observations used as an instrument. Standard errors are reported in parentheses. *, ** and *** denote significance at the 10%, 5% and 1% levels.

Columns (2) and (3) present the results of panel data regression with fixed and between effects. Both specifications indicate a selective reporting bias, which is stronger when within-study variation is considered. It is particularly pronounced when fixed effects are used, with the corrected coefficient for the size effect being reduced further to -0.0236. The next step in the analysis is estimating specification (4) using WLS. Column (4) of Table 3 reports the results with the precision variable used as a weight. As shown in Table 3, this approach produces the most pronounced publication bias and in this specification, the corrected mean size effect is reduced to a nonsignificant value. Similar, albeit less marked results are obtained in column (5), in which the baseline regression is estimated using WLS with the inverse of the number of size effect estimates reported per study as a weight. Given this and previous results, overall the funnel asymmetry test indicates a strong selective reporting bias in the estimates of the size effect.

The last part of Table 3 reports the results of instrumental variable regressions. We use the inverse of the square root of the number of observations per study as an instrument; an unreported first-step regression of the standard error on the instrumental variable produces a coefficient of 0.652 with a standard error of 0.142, which indicates the satisfactory strength of the instrument. The results of the instrumental variable regression, reported in column (6), generally confirm the preceding results of the underlying mean size effect being much smaller than the reported mean, even though the coefficient of SE is nonsignificant (the IV estimation is obviously much less precise).

Next, to investigate the pattern of publication bias, we include the interaction terms of the standard error with the recursive impact factor and standard error with the year of publication in regression (4), following Havranek and Irsova (2012), among others. The results are presented in Table 4. Below, we only present the results analogous to specifications (1) and (2) in Table 3; the results of other specifications yield a similar message and are shown in Appendix B. The findings regarding the effect of journal quality on the selective reporting bias are inconclusive, since the interaction term of the standard error and the impact factor has a negative sign in the OLS regression and a positive sign in the fixed effects regression; in both specifications, the coefficient is either nonsignificant or significant only at the 5% level. This suggests that in our dataset, journal quality is not correlated with the extent of publication bias.

Table 4: Estimating the mediating factors of publication bias

	(1) OLS	(2) OLS	(3) OLS	(4) FE	(5) FE	(6) FE
SE	-0.704***	-1.294***	-1.288***	-1.096***	-1.272***	-1.433***
	(0.175)	(0.156)	(0.128)	(0.169)	(0.178)	(0.202)
SE*Impact	-0.0686		-0.185**	0.247		0.234
	(0.0603)		(0.0766)	(0.170)		(0.159)
SE*Pub. Year		0.000164***	0.000239***		0.000135***	0.000124***
		(0.0000274)	(0.0000460)		(0.0000393)	(0.0000449)
Constant	-0.0351***	-0.0241***	-0.0247***	-0.0305**	-0.0197*	-0.0266*
	(0.00771)	(0.00857)	(0.00647)	(0.0153)	(0.0108)	(0.0141)
Observations	1663	1746	1663	1663	1746	1663

Notes: The table shows the results of regression $S_{it} = S_0 + \sigma \cdot SE(S_{it}) + \gamma \cdot SE(S_{it}) \cdot X_t + \epsilon_{it}$, where S_{it} – i-th estimate of size effect reported in study t, $SE(S_{it})$ is the standard error and X_t is either an impact factor of the outlet, in which study t was published, or the year of publication of study t. Specifications (1) – (3) are estimated using OLS with standard errors clustered by study and geographic region. Specifications (4) – (6) are panel data regressions with fixed effects. Standard errors are reported in parentheses. *, ** and *** denote significance at the 10%, 5% and 1% levels.

In contrast, the interaction term of the standard error with the year of publication is uniformly positive and highly significant, suggesting that selective reporting has been less of an issue in more recent studies on size effect, perhaps because of the use of more refined and precise econometric techniques—or a growing acceptance of positive or nonsignificant estimates following the apparent "disappearance" of the size effect in the US after the 1980s, as suggested by Horowitz et al. (2000).

Finally, we employ Hedges' model for selective reporting. We estimate the model, developed and described in detail by Hedges (1992), using four steps reflecting conventional levels of significance: p-value less than 0.01, p-value between 0.01 and 0.05, p-value between 0.05 and 0.1, and p-value more than 0.1. Hedges' model evaluates whether the weights of the estimates in the weight function (approximated by a step function with four steps described above) are different from each other; in other words, whether publication probability depends on statistical significance. The results of the estimation are reported in Table 5, where weight ω_1, associated with p-value less than 0.01, is normalized at 1. The most significant results ($p<0.01$) are the most likely to be reported, whereas the entirely nonsignificant results ($p>0.1$)

have the smallest probability of being selected for publication. The difference is statistically significant at the 0.001 level, as documented by the difference in log likelihood between the model with categories according to significance and a simple model without these categories.

Table 5: Hedges' model of selective reporting

	Unrestricted model		Restricted model	
	Coefficient	SE	Coefficient	SE
ω_2	-5.494	1.160		
ω_3	-2.723	0.953		
ω_4	-8.596	1.279		
Constant	0.015	0.004	-0.027	0.002
σ	-0.083	0.003	-0.065	0.002
Log likelihood	2669.98		2852.31	
Observations	1746		1746	
χ^2 (H_0: all estimates have the same probability of reporting): 364.66, p-value < 0.001.				

Notes: Null hypothesis of the Hedges test is the same probability of estimates significant at the 1%, 5%, and 10% levels and nonsignificant ones of being reported. ω_1, which is the weight of the probability of selection for estimates
significant at the 1% level, is normalized at 1. ω_2, ω_3 and ω_4 are the probabilities for estimates significant at the 5% level, significant at the 10% level, and nonsignificant ones are reported. σ is the standard deviation of the estimates of size effect.

We conclude the discussion of results by showing how publication bias affects the magnitude of actual size risk premium implied by the data. The size risk premium is usually estimated by sorting stocks into several portfolios formed according to the percentiles of market value of equity and then subtracting historical percentage (excess) returns on stocks in the smallest market cap portfolio from returns on stocks in the largest market cap. Such is the approach used e.g., in Grabowski et al. (2017), in which annual size premium between the 10th and 1st deciles of the CRSP US Stock Database as of December 31, 2016, is estimated at 5.94%.[3]

[3] In this study, the universe of stocks includes all stocks traded on the NYSE, NYSE-MKT (formerly AMEX) and NASDAQ, with closed-end funds, preferred stocks, REITs, foreign stocks and trusts excluded, for a

Table 6: Natural Log of Market Value of Equity Percentiles of Stocks Traded on the NYSE (in USD mil.), and implied size risk premium (in %)

Percentile	Natural log of size (USD mil.)	Annualized difference with 90th percentile (size premium), unadjusted	Annualized difference with 90th percentile (size premium), adjusted for selective reporting bias
5th	5.14	5.88%	1.98%
10th	5.83	5.08%	1.72%
15th	6.23	4.62%	1.56%
20th	6.61	4.18%	1.42%
25th	6.87	3.88%	1.31%
30th	7.13	3.59%	1.22%
35th	7.37	3.31%	1.12%
40th	7.57	3.09%	1.05%
45th	7.79	2.84%	0.97%
50th	8.01	2.59%	0.88%
55th	8.19	2.38%	0.81%
60th	8.38	2.18%	0.74%
65th	8.64	1.88%	0.64%
70th	8.90	1.58%	0.54%
75th	9.19	1.26%	0.43%
80th	9.50	0.91%	0.31%
85th	9.88	0.50%	0.17%
90th	10.33	0.00%	0.00%
95th	10.93	-0.66%	-0.23%
100th	12.92	-2.82%	-0.98%

Notes: Market value of equity breakpoints are given in millions of USD at the end of May 2017. The breakpoints use all NYSE stocks that have a CRSP share code of 10 or 11 and have good shares and price data. Closed-end funds and REITs are excluded.

To estimate size risk premium based on the reported values of size effect—that is, the slope coefficient from the regression of returns on the logarithm of the market value of equity—we use the latest available data on the size breakdown of US companies into 5th percentiles, as provided by Kenneth French and shown in Table 6.[4] Assuming a linear relationship between size and returns, the size risk premium can be calculated as the difference between the slope

total of 3,221 stocks. Data on stock prices are taken from the CRSP database. Decile breakpoints are according to the NYSE stocks market cap. Portfolios are rebalanced quarterly, and annual returns are calculated as the annualized weighted average of monthly excess returns (over CAPM) for individual stocks. The data on stock returns are for the 1926-2016 period. The size premium for individual deciles is calculated as the difference between average return in excess of CAPM and average return in excess of risk-free rate.

[4] Available at http://mba.tuck.dartmouth.edu/pages/faculty/ken.french/data_library.html. The dataset contains NYSE market equity breakpoints into 5th percentiles in million of USD; the dataset includes each company that has a CRSP share code of 10 or 11 and has good shares and price data (closed end funds and REITs excluded). The latest available data were as of May 2017.

coefficient multiplied by the 10th percentile of the market value of equity and the slope coefficient multiplied by the 90th percentile of the market value of equity.

We calculate the size risk premium unadjusted for selective reporting bias by using the simple mean of the coefficient from the regression of returns on a size of -0.092, as reported in Table 1. The difference in percentage returns between the 10th and the 90th percentile of NYSE stocks is therefore 0.415%, or 5.08% in annualized terms, consistent with the results of Grabowski et al. (2017). Additional information on the computation is shown in Table 6.

To calculate the size risk premium adjusted for the selective reporting bias, we use the estimate of -0.0315 from our baseline specification (1) reported in Table 3, which, as explained above, represents the value of the size effect corrected for bias. The implied difference in percentage returns between the 10th and the 90th percentile of NYSE stocks is 0.142%, or 1.72% in annualized terms, which is roughly three times lower than the unadjusted value.

For robustness, we repeat these computations while excluding any observations of the size effect with the returns being non-monthly, focusing on US stocks only. This reduces the number of size effect estimates to 946 observations; the mean of the slope coefficient from the regression of monthly returns on the natural log of market value of equity is -0.095, and the adjusted value of size effect from specification (1) as in Table 3 is -0.0259. This implies that the annualized size risk premium is 5.25% unadjusted, and 1.41% when adjusted for publication bias. This and previous results emphasize the general point of heavy exaggeration of size risk premium in academic research caused by, inter alia, selective reporting bias, which has direct implications for practitioners.

5. Concluding Remarks

We perform a meta-analysis of 1,746 estimates from 102 studies analyzing the impact of firm size on stock returns. We find that the literature suffers from a significant publication selection bias, which implies that the conventional estimates of the size risk premium reported in the academic literature are significantly overstated. Our estimate of the difference in annual stock returns on the smallest and largest market capitalization quintile is 1.72% after correcting for this bias, which provides a stark contrast to the uncorrected mean reported size premium of about 5%.

We use three techniques to evaluate the incidence of the publication selection bias in the literature: funnel plots, meta-regression, and Hedges' test of publication bias. The informal funnel-plot approach relies on a visual test: the true mean size premium should be close to the most precise estimates of this effect, whereas less precision brings more dispersion, which forms an inverted funnel when the size of estimates is depicted on the horizontal axis and their precision on the vertical axis in a scatter plot. Crucially, the funnel should be symmetrical in the absence of publication bias (preference for significant or negative estimates). The meta-regression technique uses the property that the ratio of the estimates and their standard errors has a t-distribution, which implies that these two should be statistically independent quantities. A regression of the reported estimates of the size effect on the reported standard error should thus bring a slope coefficient of zero if no publication bias is present. Hedges' test of publication bias examines the probability of publication for individual estimates in relation to their statistical significance.

All three methods show substantial publication selection bias: estimates that are statistically significant at conventional levels and that show the intuitive negative relationship between size and returns are much more likely to be selected for publication from the pool of all obtained results. This selective reporting has important consequences for the mean reported effect, which is exaggerated threefold, according to our results. We find that the extent of selective reporting is not related to the quality of the journal (measured by the recursive impact factor), but that it is connected to the publication year of the study: newer studies tend to be much less involved in publication selection. This marked decrease in the extent of selective reporting might have been facilitated by the increasing notion that the size premium decreased after the 1970s, which we also document in this meta-analysis. (The sudden drop occurred after the publication of the first study on the topic, Banz 1981; our finding is thus consistent with McLean and Pontiff 2016, who argue that anomalies in asset pricing become less anomalous once published.) This notion makes it easier to report nonsignificant or even positive estimates of the effect of size on returns, thereby alleviating publication bias.

References

Amihud, Yakov. 2002. "Illiquidity and Stock Returns: Cross-Section and Time-Series Effects." *Journal of Financial Markets* 5 (1), 31–56.

Ashenfelter, Orley and Michael Greenstone. 2004. *Estimating the Value of a Statistical Life: The Importance of Omitted Variables and Publication Bias*. National Bureau of Economic Research.

Asparouhova, Elena, Hendrik Bessembinder, and Ivalina Kalcheva. 2013. "Noisy Prices and Inference Regarding Returns." *The Journal of Finance* 68 (2), 665–714.

Banz, Rolf W. 1981. "The Relationship between Return and Market Value of Common Stocks." *Journal of Financial Economics* 9 (1), 3–18.

Barry, Christopher B. and Stephen J. Brown. 1984. "Differential Information and the Small Firm Effect." *Journal of Financial Economics* 13 (2), 283–94.

Basu, Sanjoy. 1983. "The Relationship between Earnings' Yield, Market Value and Return for NYSE Common Stocks : Further Evidence." *Journal of Financial Economics* 12 (1), 129–56.

Berk, Jonathan B. 2000. "Sorting out Sorts." *Journal of Finance* 55 (1), 407.

Bhandari, Laxmi Chand. 1988. "Debt/Equity Ratio and Expected Common Stock Returns: Empirical Evidence." *Journal of Finance* 43 (2), 507–28.

Black, Fischer. 1972. "Capital Market Equilibrium with Restricted Borrowing." *Journal of Business* 45 (3), 444–55.

Cakici, Nusret, Yi Tang, and An Yan. 2016. "Do the Size, Value, and Momentum Factors Drive Stock Returns in Emerging Markets?" *Journal of International Money and Finance* 69, 179–204.

Carhart, Mark M. 1997. "On Persistence in Mutual Fund Performance." *The Journal of Finance* 52 (1), 57–82.

Chan, K. C. and Nai-Fu Chen. 1991. "Structural and Return Characteristics of Small and Large Firms." *Journal of Finance* 46 (4), 1467–84.

Chen, Joseph, Harrison Hong, and Jeremy C. Stein. 2002. "Breadth of Ownership and Stock Returns." *Journal of Financial Economics* 66 (2–3), 171–205.

Chen, Nai-Fu, Richard Roll, and Stephen A. Ross. 1986. "Economic Forces and the Stock Market." *The Journal of Business* 59 (3), 383–403.

Chordia, Tarun and Lakshmanan Shivakumar. 2002. "Momentum, Business Cycle, and Time-Varying Expected Returns." *Journal of Finance* 57 (2), 985.

Conrad, Jennifer and Gautam Kaul. 1998. "An Anatomy of Trading Strategies." *Review of Financial Studies* 11 (3), 489.

Crain, Michael A. 2011. *A Literature Review of the Size Effect*. Rochester, NY: Social Science Research Network.

Dichev, Ilia D. 1998. "Is the Risk of Bankruptcy a Systematic Risk?" *The Journal of Finance* 53 (3), 1131–47.

van Dijk, Mathijs A. 2011. "Is Size Dead? A Review of the Size Effect in Equity Returns." *Journal of Banking & Finance* 35 (12), 3263–74.

Dissanaike, Gishan. 2002. "Does the Size Effect Explain the UK Winner-Loser Effect?" *Journal of Business Finance & Accounting* 29 (1–2), 139–154.

Doucouliagos, Chris and Tom D. Stanley. 2013. “Are All Economic Facts Greatly Exaggerated? Theory Competition and Selectivity.” *Journal of Economic Surveys* 27 (2), 316–339.

Easterday, Kathryn E., Pradyot K. Sen, and Jens A. Stephan. 2009. “The Persistence of the Small Firm/January Effect: Is It Consistent with Investors’ Learning and Arbitrage Efforts?” *The Quarterly Review of Economics and Finance* 49 (3), 1172–93.

Egger, Matthias, George Davey Smith, Martin Schneider, and Christoph Minder. 1997. “Bias in Meta-Analysis Detected by a Simple, Graphical Test.” *Bmj* 315 (7109), 629–634.

Fama, Eugene F. and Kenneth R. French. 1992. “The Cross-Section of Expected Stock Returns.” *The Journal of Finance* 47 (2), 427–65.

Fama, Eugene F. and Kenneth R. French. 1993. “Common Risk Factors in the Returns on Stocks and Bonds.” *Journal of Financial Economics* 33 (1), 3–56.

Fama, Eugene F. and Kenneth R. French. 2008. “Dissecting Anomalies.” *The Journal of Finance* 63 (4), 1653–78.

Fama, Eugene F. and Kenneth R. French. 2015. “A Five-Factor Asset Pricing Model.” *Journal of Financial Economics* 116 (1), 1–22.

Fama, Eugene F. and James D. MacBeth. 1973. “Risk, Return, and Equilibrium: Empirical Tests.” *The journal of political economy* 607–636.

Feld, Lars, Jost Heckemeyer, and Michael Overesch. 2013. “Capital structure choice and company taxation: A meta-study.” *Journal of Banking and Finance 37(8), 2850-2866.*

Feld, Lars and Jost Heckemeyer. 2011. “FDI And Taxation: A Meta-Study.” *Journal of Economic Surveys 25(2), 233-272.*

Fletcher, Jonathan. 1997. “An Examination of the Cross-Sectional Relationship of Beta and Return: UK Evidence.” *Journal of Economics and Business* 49 (3), 211–221.

Grabowski, Roger J., James P. Harrington, and Carla Nunes. 2017. *Valuation Handbook-US Guide to Cost of Capital*. John Wiley & Sons.

Graham, John R. and Campbell R. Harvey. 2001. “The Theory and Practice of Corporate Finance: Evidence from the Field.” *Journal of Financial Economics* 60 (2–3), 187–243.

Harvey, Campbell R., Yan Liu, and Heqing Zhu. 2016. “… and the Cross-Section of Expected Returns.” *Review of Financial Studies* 29 (1), 5–68.

Havranek, Tomas. 2015 “Measuring Intertemporal Substitution: The Importance Of Method Choices And Selective Reporting.” *Journal of the European Economic Association* 13 (6), 1180-1204.

Havranek, Tomas, Roman Horvath, Zuzana Irsova, and Marek Rusnak. 2015. " Cross-country heterogeneity in intertemporal substitution." *Journal of International Economics* 96, 100-118.

Havranek, Tomas and Zuzana Irsova. 2012. “Survey Article: Publication Bias in the Literature on Foreign Direct Investment Spillovers.” *Journal of Development Studies* 48 (10), 1375-1396.

Havranek, Tomas and Zuzana Irsova. 2011. “Estimating Vertical Spillovers from FDI: Why Results Vary and What the True Effect Is.” *Journal of International Economics* 85 (2), 234–44.

Havranek, Tomas and Zuzana Irsova. 2017. “Do Borders Really Slash Trade? A Meta-Analysis.” *IMF Economic Review* 65 (2), 365–96.

Havranek, Tomas, Marek Rusnak, and Anna Sokolova. 2017. "Habit Formation in Consumption: A Meta-Analysis." *European Economic Review* 95, 142-167.

Hedges, Larry V. 1992. “Modeling Publication Selection Effects in Meta-Analysis.” *Statistical Science* 246–255.

Horowitz, Joel L., Tim Loughran, and N. E. Savin. 2000. “The Disappearing Size Effect.” *Research in Economics* 54 (1), 83–100.

Horowitz, Joel L., Tim Loughran, and N. E. Savin. 2000. “Three Analyses of the Firm Size Premium.” *Journal of Empirical Finance* 7 (2), 143–53.

Hou, Kewei, G.Andrew Karolyi, and Bong-Chan Kho. 2011. “What Factors Drive Global Stock Returns?” *Review of Financial Studies* 24 (8), 2527–74.

Hur, Jungshik, Glenn Pettengill, and Vivek Singh. 2014. “Market States and the Risk-Based Explanation of the Size Premium.” *Journal of Empirical Finance* 28, 139–50.

Jegadeesh, Narasimhan and Sheridan Titman. 1993. “Returns to Buying Winners and Selling Losers: Implications for Stock Market Efficiency.” *Journal of Finance* 48 (1), 65–91.

Keim, Donald B. 1983. “Size-Related Anomalies and Stock Return Seasonality.” *Journal of Financial Economics* 12 (1), 13–32.

Knez, Peter J. and Mark J. Ready. 1997. “On The Robustness of Size and Book-to-Market in Cross-Sectional Regressions.” *The Journal of Finance* 52 (4), 1355–82.

Lakonishok, Josef, Andrei Shleifer, and Robert W. Vishny. 1994. “Contrarian Investment, Extrapolation, and Risk.” *The journal of finance* 49 (5), 1541–1578.

Lee, C. M. C. 2001. “Market Efficiency and Accounting Research: A Discussion Of’capital Market Research in Accounting’by SP Kothari.” *Journal of Accounting and Economics* 31 (1), 233–253.

Lintner, John. 1965. “The Valuation of Risk Assets and the Selection of Risky Investments in Stock Portfolios and Capital Budgets.” *Review of Economics & Statistics* 47 (1), 13.

Lo, Andrew W. and A.Craig MacKinlay. 1990. “Data-Snooping Biases in Tests of Financial Asset Pricing Models.” *The Review of Financial Studies* 3 (3), 431–67.

MacKinlay, A.Craig. 1995. “Multifactor Models Do Not Explain Deviations from the CAPM.” *Journal of Financial Economics* 38 (1), 3–28.

McLean, R. David and Jeffrey Pontiff. 2016. “Does Academic Research Destroy Stock Return Predictability?” *Journal of Finance* 71 (1), 5–32.

Reinganum, Marc R. 1981. “Misspecification of Capital Asset Pricing: Empirical Anomalies Based on Earnings’ Yields and Market Values.” *Journal of Financial Economics* 9 (1), 19–46.

Rusnak, Marek, Tomas Havranek, and Roman Horvath. 2013. “How to Solve the Price Puzzle? A Meta-Analysis.” *Journal of Money, Credit and Banking* 45 (1), 37–70.

Sharpe, William F. 1964. “Capital Asset Prices: A Theory of Market Equilibrium under Conditions of Risk.” *Journal of Finance* 19 (3), 425.

Stanley, Tom D. 2005. “Beyond Publication Bias.” *Journal of Economic Surveys* 19(3), 309-345.

Stanley, Tom D. 2008. “Meta-Regression Methods for Detecting and Estimating Empirical Effects in the Presence of Publication Selection.” *Oxford Bulletin of Economics and statistics* 70 (1), 103–127.

Vassalou, Maria and Yuhang Xing. 2004. “Default Risk in Equity Returns.” *The Journal of Finance* 59 (2), 831–68.

Zigraiova, D. and T. Havranek. 2016. “Bank Competition And Financial Stability: Much Ado About Nothing?” *Journal of Economic Surveys* 30(5), 944-981

Appendix A – List of Studies Used in the Meta-Analysis

Study	# of citations	Recursive impact factor
Acharya, Viral V. and Lasse Heje Pedersen. 2005. "Asset Pricing with Liquidity Risk." Journal of financial Economics 77 (2), 375–410.	2676	1.498
Amel-Zadeh, Amir. 2011. "The Return of the Size Anomaly: Evidence from the German Stock Market." European Financial Management 17 (1), 145–182.	30	0.162
Amihud, Yakov. 2002. "Illiquidity and Stock Returns: Cross-Section and Time-Series Effects." Journal of Financial Markets 5 (1), 31–56.	558	1.776
Amihud, Yakov and Haim Mendelson. 1989. "The Effects of Beta, Bid-Ask Spread, Residual Risk, and Size on Stock Returns." The Journal of Finance 44 (2), 479–486.	4956	0.387
Anderson, Christopher W. and Luis Garcia-Feijóo. 2006. "Empirical Evidence on Capital Investment, Growth Options, and Security Returns." The Journal of Finance 61 (1), 171–94.	286	1.776
Ang, Andrew, Joseph Chen, and Yuhang Xing. 2006. "Downside Risk." Review of Financial Studies 19 (4), 1191–1239.	859	1.498
Ang, Andrew, Robert J. Hodrick, Yuhang Xing, and Xiaoyan Zhang. 2009. "High Idiosyncratic Volatility and Low Returns: International and Further U.S. Evidence." Journal of Financial Economics 91 (1), 1–23.	637	2.081
Asparouhova, Elena, Hendrik Bessembinder, and Ivalina Kalcheva. 2013. "Noisy Prices and Inference Regarding Returns." The Journal of Finance 68 (2), 665–714.	77	1.776
Avramov, Doron and Tarun Chordia. 2006. "Asset Pricing Models and Financial Market Anomalies." Review of Financial Studies 19 (3), 1001–1040.	369	2.081
Bagella, Michele, Leonardo Becchetti, and Andrea Carpentieri. 2000. "'The First Shall Be Last'. Size and Value Strategy Premia at the London Stock Exchange." Journal of banking & finance 24 (6), 893–919.	32	0.208
Bali, Turan G., Nusret Cakici, and Robert F. Whitelaw. 2011. "Maxing out: Stocks as Lotteries and the Cross-Section of Expected Returns." Journal of Financial Economics 99 (2), 427–446.	339	1.498
Banz, Rolf W. 1981. "The Relationship between Return and Market Value of Common Stocks." Journal of financial economics 9 (1), 3–18.	5866	1.498
Barry, Christopher B. and Stephen J. Brown. 1984. "Differential Information and the Small Firm Effect." Journal of Financial Economics 13 (2), 283–294.	602	1.498
Barry, Christopher B., Elizabeth Goldreyer, Larry Lockwood, and Mauricio Rodriguez. 2002. "Robustness of Size and Value Effects in Emerging Equity Markets, 1985–2000." Emerging Markets Review 3 (1), 1–30.	125	0.094
Bauer, Rob, Mathijs Cosemans, and Peter C. Schotman. 2010. "Conditional Asset Pricing and Stock Market Anomalies in Europe." European Financial Management 16 (2), 165–190.	52	0.162
Bhandari, Laxmi Chand. 1988. "Debt/Equity Ratio and Expected Common Stock Returns: Empirical Evidence." The journal of finance 43 (2), 507–528.	1308	1.776
Borys, Magdalena Morgese and Petr Zemčik. 2011. "Size and Value Effects in the Visegrad Countries." Emerging Markets Finance and Trade 47 (3), 50–68.	15	0.015

Study	# of citations	Recursive impact factor
Brammer, Stephen, Chris Brooks, and Stephen Pavelin. 2006. "Corporate Social Performance and Stock Returns: UK Evidence from Disaggregate Measures." Financial management 35 (3), 97–116.	552	0.236
Brennan, Michael, Tarun Chordia, and Avanidhar Subrahmanyam. 2004. "Cross-Sectional Determinants of Expected Returns." The Legacy of Fischer Black 161.	66	1.498
Brennan, Michael J., Tarun Chordia, and Avanidhar Subrahmanyam. 1998. "Alternative Factor Specifications, Security Characteristics, and the Cross-Section of Expected Stock Returns." Journal of Financial Economics 49 (3), 345–373.	1335	1.498
Brennan, Michael J., Tarun Chordia, Avanidhar Subrahmanyam, and Qing Tong. 2012. "Sell-Order Liquidity and the Cross-Section of Expected Stock Returns." Journal of Financial Economics 105 (3), 523–541.	37	n/a
Bryant, Paul S. and Venkat R. Eleswarapu. 1997. "Cross-Sectional Determinants of New Zealand Share Market Returns." Accounting & Finance 37 (2), 181–205.	38	n/a
Burlacu, Radu, Patrice Fontaine, Sonia Jimenez-Garces, and Mark S. Seasholes. 2012. "Risk and the Cross Section of Stock Returns." Journal of financial economics 105 (3), 511–522.	10	1.498
Chan, Andrew and Alice P. L. Chui. 1996. "An Empirical Re-Examination of the Cross-Section of Expected Returns: UK Evidence." Journal of Business Finance & Accounting 23 (9–10), 1435–52.	177	n/a
Chan, K. C. and Nai-Fu Chen. 1988. "An Unconditional Asset-Pricing Test and the Role of Firm Size as an Instrumental Variable for Risk." The Journal of Finance 43 (2), 309–325.	409	1.776
Chan, K. C. and Nai-Fu Chen. 1991. "Structural and Return Characteristics of Small and Large Firms." The Journal of Finance 46 (4), 1467–1484.	671	1.776
Chan, Kevin C., Nai-fu Chen, and David A. Hsieh. 1985. "An Exploratory Investigation of the Firm Size Effect." Journal of Financial Economics 14 (3), 451–471.	180	2.081
Chan, Louis K. C., Yasushi Hamao, and Josef Lakonishok. 1991. "Fundamentals and Stock Returns in Japan." The Journal of Finance 46 (5), 1739–64.	90	1.776
Chan, Louis K. C., Yasushi Hamao, and Josef Lakonishok. 1993. "Can Fundamentals Predict Japanese Stock Returns?" Financial Analysts Journal 49 (4), 63–69.	1187	1.776
Chan, Louis K. C., Jason Karceski, and Josef Lakonishok. 1998. "The Risk and Return from Factors." Journal of Financial and Quantitative Analysis.	68	0.027
Chang, Rosita, Liming Guan, Jianguo Chen, Kwong Leong Kan, and Hamish Anderson. 2007. "Size, Book/Market Ratio and Risk Factor Returns: Evidence from China A-Share Market." Managerial Finance 33 (8), 574–594.	1466	1.776
Chen, Joseph, Harrison Hong, and Jeremy C. Stein. 2002. "Breadth of Ownership and Stock Returns." Journal of Financial Economics 66 (2–3), 171–205.	903	1.776
Chopra, Navin, Josef Lakonishok, and Jay R. Ritter. 1992. "Measuring Abnormal Performance." Journal of Financial Economics 31 (2), 235–68.	48	0.044
Chordia, Tarun, Avanidhar Subrahmanyam, and V.Ravi Anshuman. 2001. "Trading Activity and Expected Stock Returns." Journal of Financial Economics 59 (1), 3–32.	23	0.074

Study	# of citations	Recursive impact factor
Chui, Andy C. W. and K. C.John Wei. 1998. "Book-to-Market, Firm Size, and the Turn-of-the-Year Effect: Evidence from Pacific-Basin Emerging Markets." Pacific-Basin Finance Journal 6 (3–4), 275–93.	1828	1.776
Claessens, Stijn, Susmita Dasgupta, and Jack Glen. 1995. The Cross-Section of Stock Returns: Evidence from the Emerging Markets. World Bank Publications.	346	1.498
Cooper, Michael J., Huseyin Gulen, and Michael J. Schill. 2008. "Asset Growth and the Cross-Section of Stock Returns." The Journal of Finance 63 (4), 1609–51.	256	1.776
Cooper, Michael J., Huseyin Gulen, and Alexei V. Ovtchinnikov. 2010. "Corporate Political Contributions and Stock Returns." The Journal of Finance 65 (2), 687–724.	45	0.096
Cremers, KJ Martijn, Vinay B. Nair, and Kose John. 2009. "Takeovers and the Cross-Section of Returns." Review of Financial Studies 22 (4), 1409–1445.	59	0.032
Da, Zhi. 2009. "Cash Flow, Consumption Risk, and the Cross-Section of Stock Returns." The Journal of Finance 64 (2), 923–956.	1102	1.776
Datar, Vinay T., Narayan Y. Naik, and Robert Radcliffe. 1998. "Liquidity and Stock Returns: An Alternative Test." Journal of Financial Markets 1 (2), 203–19.	15628	1.776
Diavatopoulos, Dean, James S. Doran, and David R. Peterson. 2008. "The Information Content in Implied Idiosyncratic Volatility and the Cross-Section of Stock Returns: Evidence from the Option Markets." Journal of Futures Markets 28 (11), 1013–1039.	5	0.096
Dichev, Ilia D. 1998. "Is the Risk of Bankruptcy a Systematic Risk?" The Journal of Finance 53 (3), 1131–47.	819	1.776
Diether, Karl B., Christopher J. Malloy, and Anna Scherbina. 2002. "Differences of Opinion and the Cross Section of Stock Returns." The Journal of Finance 57 (5), 2113–41.	132	0.082
Dissanaike, Gishan. 2002. "Does the Size Effect Explain the UK Winner-Loser Effect?" Journal of Business Finance & Accounting 29 (1–2), 139–154.	702	1.498
Doeswijk, Ronald Q. 1997. "Contrarian Investment in the Dutch Stock Market." De Economist 145 (4), 573–598.	40	n/a
Easley, David, Soeren Hvidkjaer, and Maureen O'Hara. 2002. "Is Information Risk a Determinant of Asset Returns?" The Journal of Finance 57 (5), 2185–2221.	61	0.014
Eleswarapu, Venkat R. 1997. "Cost of Transacting and Expected Returns in the Nasdaq Market." The Journal of Finance 52 (5), 2113–27.	233	1.498
Eleswarapu, Venkat R. and Marc R. Reinganum. 1993. "The Seasonal Behavior of the Liquidity Premium in Asset Pricing." Journal of Financial Economics 34 (3), 373–386.	73	0.208
Elfakhani, Said, Larry J. Lockwood, and Tarek S. Zaher. 1998. "Small Firm and Value Effects in the Canadian Stock Market." Journal of Financial Research 21 (3), 277–91.	150	0.162
Estrada, Javier and Ana Paula Serra. 2005. "Risk and Return in Emerging Markets: Family Matters." Journal of Multinational Financial Management 15 (3), 257–272.	24	0.067
Fama, Eugene F. and Kenneth R. French. 1992. "The Cross-Section of Expected Stock Returns." the Journal of Finance 47 (2), 427–465.	92	0.082
Fama, Eugene F. and Kenneth R. French. 2008. "Dissecting Anomalies." The Journal of Finance 63 (4), 1653–78.	116	0.128

Study	# of citations	Recursive impact factor
Fan, Xinting and Ming Liu. 2005. "Understanding Size and the Book-to-market Ratio: An Empricial Exploration of Berk's Critique." Journal of Financial Research 28 (4), 503–518.	126	0.288
Ferson, Wayne E. and Campbell R. Harvey. 1999. "Conditioning Variables and the Cross Section of Stock Returns." The Journal of Finance 54 (4), 1325–60.	473	2.081
Fletcher, Jonathan. 1997. "An Examination of the Cross-Sectional Relationship of Beta and Return: UK Evidence." Journal of Economics and Business 49 (3), 211–221.	260	2.081
Fu, Fangjian. 2009. "Idiosyncratic Risk and the Cross-Section of Expected Stock Returns." Journal of Financial Economics 91 (1), 24–37.	174	0.381
Garza-Gómez, Xavier, Jiro Hodoshima, and Michio Kunimura. 1998. "Does Size Really Matter in Japan?" Financial Analysts Journal 54 (6), 22–34.	84	0.046
Gaunt, Clive, Philip Gray, and Julie McIvor. 2000. "The Impact of Share Price on Seasonality and Size Anomalies in Australian Equity Returns." Accounting & Finance 40 (1), 33–50.	1	0.288
George, Thomas J. and Chuan-Yang Hwang. 2010. "A Resolution of the Distress Risk and Leverage Puzzles in the Cross Section of Stock Returns." Journal of Financial Economics 96 (1), 56–79.	335	1.776
Herrera, Martin J. and Larry J. Lockwood. 1994. "The Size Effect in the Mexican Stock Market." Journal of Banking & Finance 18 (4), 621–32.	871	1.776
Heston, Steven L., K.Geert Rouwenhorst, and Roberto E. Wessels. 1999. "The Role of Beta and Size in the Cross-Section of European Stock Returns." European Financial Management 5 (1), 9–27.	53	0.044
Ho, Ron Yiu-wah, Roger Strange, and Jenifer Piesse. 2006. "On the Conditional Pricing Effects of Beta, Size, and Book-to-Market Equity in the Hong Kong Market." Journal of International Financial Markets, Institutions and Money 16 (3), 199–214.	65	n/a
Hodoshima, Jiro, Xavier Garza–Gómez, and Michio Kunimura. 2000. "Cross-Sectional Regression Analysis of Return and Beta in Japan." Journal of Economics and Business 52 (6), 515–33.	1732	1.776
Horowitz, Joel L., Tim Loughran, and N. E. Savin. 2000. "The Disappearing Size Effect." Research in Economics 54 (1), 83–100.	630	1.498
Horowitz, Joel L., Tim Loughran, and N. E. Savin. 2000. "Three Analyses of the Firm Size Premium." Journal of Empirical Finance 7 (2), 143–53.	393	0.38
Hou, Kewei and Tobias J. Moskowitz. 2005. "Market Frictions, Price Delay, and the Cross-Section of Expected Returns." Review of Financial Studies 18 (3), 981–1020.	32	0.008
Hou, Kewei, G.Andrew Karolyi, and Bong-Chan Kho. 2011. "What Factors Drive Global Stock Returns?" Review of Financial Studies 24 (8), 2527–2574.	1001	1.498
Hou, Kewei, Mathijs A. van Dijk, and Yinglei Zhang. 2012. "The Implied Cost of Capital: A New Approach." Journal of Accounting and Economics 53 (3), 504–26.	1008	1.498
Howton, Shelly W. and David R. Peterson. 1998. "An Examination of Cross-Sectional Realized Stock Returns Using a Varying-Risk Beta Model." Financial Review 33 (3), 199–212.	706	1.498

Study	# of citations	Recursive impact factor
Hur, Jungshik, Glenn Pettengill, and Vivek Singh. 2014. “Market States and the Risk-Based Explanation of the Size Premium.” Journal of Empirical Finance 28, 139–50.	265	0.057
Jaffe, Jeffrey, Donald B. Keim, and Randolph Westerfield. 1989. “Earnings Yields, Market Values, and Stock Returns.” The Journal of Finance 44 (1), 135–148.	561	1.776
Jagannathan, Ravi and Yong Wang. 2007. “Lazy Investors, Discretionary Consumption, and the Cross-Section of Stock Returns.” The Journal of Finance 62 (4), 1623–1661.	207	1.776
Jagannathan, Ravi and Zhenyu Wang. 1996. “The Conditional CAPM and the Cross-Section of Expected Returns.” The Journal of Finance 51 (1), 3–53.	2424	1.776
Jegadeesh, Narasimhan. 1992. “Does Market Risk Really Explain the Size Effect?” The Journal of Financial and Quantitative Analysis 27 (3), 337–51.	172	0.38
Jensen, Gerald R. and Jeffrey M. Mercer. 2002. “Monetary Policy and the Cross-Section of Expected Stock Returns.” Journal of Financial Research 25 (1), 125–39.	82	0.096
Kim, Dongcheol. 1997. “A Reexamination of Firm Size, Book-to-Market, and Earnings Price in the Cross-Section of Expected Stock Returns.” Journal of Financial and Quantitative Analysis 32 (04), 463–489.	135	0.38
Knez, Peter J. and Mark J. Ready. 1997. “On The Robustness of Size and Book-to-Market in Cross-Sectional Regressions.” The Journal of Finance 52 (4), 1355–82.	340	1.776
Kothari, S. P., Jay Shanken, and Richard G. Sloan. 1995. “Another Look at the Cross-Section of Expected Stock Returns.” The Journal of Finance 50 (1), 185–224.	1279	1.776
La Porta, Rafael. 1996. “Expectations and the Cross-Section of Stock Returns.” The Journal of Finance 51 (5), 1715–42.	893	1.776
Lakonishok, Josef and Alan C. Shapiro. 1986. “Systematic Risk, Total Risk and Size as Determinants of Stock Market Returns.” Journal of Banking & Finance 10 (1), 115–132.	342	0.208
Lakonishok, Josef, Andrei Shleifer, and Robert W. Vishny. 1994. “Contrarian Investment, Extrapolation, and Risk.” The Journal of Finance 49 (5), 1541–78.	4840	1.776
Loughran, Tim. 1997. “Book-to-Market across Firm Size, Exchange, and Seasonality: Is There an Effect?” Journal of financial and quantitative analysis 32 (03), 249–268.	3964	1.776
Loughran, Tim and Jay R. Ritter. 1995. “The New Issues Puzzle.” The Journal of Finance 50 (1), 23–51.	401	0.38
Moskowitz, Tobias J. and Mark Grinblatt. 1999. “Do Industries Explain Momentum?” The Journal of Finance 54 (4), 1249–90.	1418	1.776
Nagel, Stefan. 2005. “Short Sales, Institutional Investors and the Cross-Section of Stock Returns.” Journal of Financial Economics 78 (2), 277–309.	678	1.498
Novy-Marx, Robert. 2013. “The Other Side of Value: The Gross Profitability Premium.” Journal of Financial Economics 108 (1), 1–28.	351	1.498
Ozgur Demirtas, K. and A. Burak Guner. 2008. “Can Overreaction Explain Part of the Size Premium?” International Journal of Revenue Management 2 (3), 234–253.	5	0.002

Study	# of citations	Recursive impact factor
Penman, Stephen H., Scott A. Richardson, and İrem Tuna. 2007. "The Book-to-Price Effect in Stock Returns: Accounting for Leverage." Journal of Accounting Research 45 (2), 427–67.	209	0.304
Pettengill, Glenn, Sridhar Sundaram, and Ike Mathur. 2002. "Payment For Risk: Constant Beta Vs. Dual-Beta Models." Financial Review 37 (2), 123–135.	55	0.046
Phalippou, Ludovic. 2007. "Can Risk-Based Theories Explain the Value Premium?" Review of Finance 11 (2), 143–66.	52	0.34
Pinfold, John F., William R. Wilson, and Qiuli Li. 2001. "Book-to-Market and Size as Determinants of Returns in Small Illiquid Markets: The New Zealand Case." Financial Services Review 10 (1–4), 291–302.	23	0.034
Pontiff, Jeffrey and Artemiza Woodgate. 2008. "Share Issuance and Cross-Sectional Returns." The Journal of Finance 63 (2), 921–45.	381	1.776
Reinganum, Marc R. 1982. "A Direct Test of Roll's Conjecture on the Firm Size Effect." The Journal of Finance 37 (1), 27–35.	270	1.776
Serra, Ana Paula. 2003. "The Cross-Sectional Determinants of Returns: Evidence from Emerging Markets' Stocks." Journal of Emerging Market Finance 2 (2), 123–162.	31	0.021
Strong, Norman and Xinzhong G. Xu. 1997. "Explaining the Cross-Section of UK Expected Stock Returns." The British Accounting Review 29 (1), 1–23.	174	n/a
Tinic, Seha M. and Richard R. West. 1986. "Risk, Return, and Equilibrium: A Revisit." Journal of Political Economy 94 (1), 126–47.	212	4.107
Vos, Ed and Byron Pepper. 1997. "The Size and Book to Market Effects in New Zealand." New Zealand Investment Analyst 18, 35–45.	21	n/a
Wahlroos, Björn. 1986. "Anomalies and Equilibrium Returns in a Small Stock Market." Journal of Business Research 14 (5), 423–40.	16	0.007
Waszczuk, A. 2013. "A Risk-Based Explanation of Return Patterns-Evidence from the Polish Stock Market." Emerging Markets Review 15, 186–210.	39	0.094
Whited, Toni M. and Guojun Wu. 2006. "Financial Constraints Risk." Review of Financial Studies 19 (2), 531–59.	981	2.081
Zarowin, Paul. 1990. "Size, Seasonality, and Stock Market Overreaction." Journal of Financial and Quantitative Analysis 25 (01), 113–125.	459	0.38

Appendix B – Estimating the mediating factors of publication bias: additional results

	(1) BE	(2) BE	(3) BE	(4) WLS - Precision	(5) WLS - Precision	(6) WLS - Precision	(7) WLS – 1/n	(8) WLS – 1/n	(9) WLS – 1/n
SE	0.108	-0.741***	-0.670***	-1.180***	-1.732***	-1.749***	-0.0259	-0.835	-0.526***
	(0.115)	(0.206)	(0.167)	(0.0653)	(0.0663)	(0.0748)	(0.118)	(0.526)	(0.152)
SE*Impact	-0.210***		-0.348***	0.00863		-0.101***	-0.177***		-0.263***
	(0.0334)		(0.0373)	(0.0287)		(0.0309)	(0.0320)		(0.0550)
SE*Pub. Year		0.0000547	0.000252***		0.000232***	0.000276***		0.0000711	0.000158***
		(0.0000456)	(0.0000435)		(0.0000141)	(0.0000452)		(0.000102)	(0.0000292)
Constant	-0.0562***	-0.0347	-0.0396**	-0.000280	-0.000127	-0.0000701	-0.0490***	-0.0304	-0.0377***
	(0.0202)	(0.0247)	(0.0176)	(0.000275)	(0.000246)	(0.000244)	(0.00824)	(0.0195)	(0.00883)
Observations	1663	1746	1663	1663	1746	1663	1663	1746	1663

Notes: The table shows the results of regression $S_{it} = S_0 + \sigma \cdot SE(S_{it}) + \gamma \cdot SE(S_{it}) \cdot X_t + \epsilon_{it}$, where S_{it} – i-th estimate of size effect reported in study t, $SE(S_{it})$ is the standard error and X_t is either an impact factor of the outlet in which study t was published or the year of publication of study t. Specifications (1)-(3) are panel data regressions with between effects. Specifications (4)-(6) are estimated using WLS with precision as a weight. Specifications (7)-(9) are estimated using WLS with the reciprocal of the number of size effect estimates reported per study as a weight. Specification (6) is the instrumental variables regression with the reciprocal of the square root of the number of observations used as an instrument, with standard errors clustered by study and geographic region. Standard errors are reported in parentheses. *, ** and *** denote significance at the 10%, 5% and 1% levels.