

Firm Size and Stock Returns: A Quantitative Survey

Anton Astakhov, Tomas Havranek, and Jiri Novak

Institute of Economic Studies, Faculty of Social Sciences, Charles University, Prague

Abstract

Firm size is commonly used in numerous empirical asset pricing models as a determinant of expected stock returns. Yet there is little consensus over the magnitude and stability of the size premium. In fact, some researchers even question whether firm size should be used as a pricing factor. We collect 1,746 estimates of the slope coefficients capturing the association between firm size and stock returns reported in 102 published studies and conduct the first meta-analysis on the size premium. We find evidence of a strong bias towards publishing statistically significant negative slope coefficients. After correcting for the bias, we find that the literature implies a difference in annual stock returns on the smallest and the largest New York Stock Exchange (NYSE) market capitalization quintiles of 1.72 per cent. For the time periods covered in the sampled articles we find that the size premium was larger in earlier years and 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

C18, G12, G15

The data and codes used in this study are available in an online appendix at <http://meta-analysis.cz/size>. Corresponding author: Tomas Havranek, tomas.havranek@ies-prague.org. The authors acknowledge support from the Czech Science Foundation (project #18-02513S). Havranek also acknowledges support from Charles University (project Primus/17/HUM/16).

1. Introduction

The identification of factors explaining variation in stock returns and estimating the magnitude of the corresponding 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 their systematic risk. The Capital Asset Pricing Model (CAPM; Sharpe 1964; Lintner 1965; Black 1972) postulates that an asset's systematic risk depends on its contribution to the volatility of returns on the market portfolio. The model predicts a positive linear association between an asset's expected return and its systematic risk defined as the sensitivity of the asset's returns to the variation in market returns. CAPM also suggests that the market beta is a comprehensive 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 the bankruptcy, liquidity, and information risks.

We perform a meta-analysis of published studies on the association between firm size and realized stock returns. Firm size is one of the first empirically documented firm characteristics associated with realized stock returns (Banz 1981; Reinganum 1981; Keim 1983). Fama and French (1992) consider the size effect "the most prominent" empirical

contradiction to the CAPM. Firm size is used as a factor in all major empirical asset pricing models, i.e. the three-factor model (Fama and French 1993), the four-factor model (Carhart 1997), and the five-factor model (Fama and French 2015). Indeed, Chen et al. (1986) argue that *“size may be the best theory we now have of expected returns”* (p. 394).

Despite its widespread use as a pricing factor, firm size exhibits several puzzling features that are difficult to reconcile with the notion of a risk proxy. The size premium (i.e. the difference between average stock returns earned on small and large stocks) is not consistent across regions (Hou et al. 2011; Cakici et al. 2016) and across time (van Dijk 2011; Joel L. Horowitz et al. 2000b). It is concentrated in the month of January (Keim 1983) and in the very smallest stocks (Banz 1981; Knez and Ready 1997). Contrary to the prediction that risky assets yield low returns in times of economic declines, the size premium is larger in “bear” markets than in “bull” markets (Hur, Pettengill, and Singh 2014). Furthermore, it is not obvious why large stocks should be considered less risky while stocks that have recently grown in terms of their market capitalization, i.e. stocks with a positive stock price momentum, shall be more risky (Conrad and Kaul 1998; Chordia and Shivakumar 2002; Carhart 1997).

These puzzling findings fuel the controversy over whether 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 followed by a stock price correction, 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 empirical findings. Thus, it is important to systematically study the evidence on the association between firm size and stock returns and to evaluate its strength and consistency.

Unsurprisingly, prior research has amassed an extensive body of empirical evidence on the explanatory power of firm size in various settings (for an overview see Harvey et al. 2016). Performing a meta-analysis allows us to systematically aggregate and synthesize past findings, to evaluate the strength of the association between firm size and stock returns, and to assess its stability over time. Furthermore, this methodological approach allows us to adjust for a potential publication selection bias among the reported estimates and to estimate the magnitude of the size premium avoiding the criticized pre-sorting of stocks into portfolios based on previously empirically documented characteristics. Our study is thus relevant both for the development of finance theory that aims at explaining the existence of empirical pricing factors, as well as for practical finance application, including the estimation of the cost of capital and the performance evaluation of investment strategies.

We document significant publication selection bias in the existing literature. In line with prior literature we define the size premium as difference in stock returns on the smallest and largest market capitalization portfolios of stocks traded on the NYSE, NYSE-MKT (formerly AMEX) and NASDAQ using NYSE breakpoints published on Kenneth French's website¹. After adjusting for the effect of publication selection we estimate the size premium 1.72%. We assume a linear relationship between firm size and stock returns, which implies that our size premium estimate is applicable for firms of sizes comparable with the NYSE market capitalization quintiles and it is not necessarily generalizable to micro stocks at the bottom or below the smallest NYSE quintile. In addition, as most of the studies that constitute our sample are based on the U.S. stock market data our size premium estimate applies to comparable markets characterized by an effective institutional regulatory framework and high liquidity. We

¹ http://mba.tuck.dartmouth.edu/pages/faculty/ken.french/data_library.html.

observe a large regional variation in the reported slope coefficients that suggest that the premium is likely larger in less well-regulated markets. Furthermore, we observe larger coefficients in the month of January and slightly larger coefficients in studies that do not control for the market beta. Surprisingly, the quality of the journal does not seem to affect the magnitude of the publication bias.

We also find that the size premium was larger prior to the publication of the first study on the topic and that newer studies are less affected by selective publication. A possible explanation for the decrease in the publication bias over time might be the growing prevalence of the view that the magnitude of the size premium decreased after the effect was first identified by Banz (1981). We show a drop of approximately 50 per cent in the estimated magnitude of the size premium in studies conducted on observations past 1981 compared to studies conducted on observations from earlier years. The growing acceptance of the smaller size premium makes it easier for newer studies to report nonsignificant or positive slope coefficients. However, we acknowledge that our analysis is constrained by the sample periods covered in the articles that we include in our study and that our approach allows us to detect potential recent shifts in the size premium only after the new data are included in samples of newly published academic studies. Future research is needed to analyze whether the decline in the size premium after 1980s has been recently reversed.

The remainder of the paper is organized as follows. Section 2 reviews the related literature, Section 3 discusses methodology and data, 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; Appendix C shows the results after including data from working papers.

2. Related Literature

The negative association between firm size (i.e. its market capitalization) and realized stock returns was first documented in early 1980s (Banz 1981; Reinganum 1981; Keim 1983). The size effect, i.e. the tendency of small firms' stocks to earn higher returns than large firms' stocks, was puzzling because the relationship had not been envisaged by asset pricing theory, such as the CAPM. The anomalous nature of this finding motivated extensive follow up empirical research aimed at analyzing the size effect in various contexts. This stream of research documented the size premium in various markets and time periods. The fairly high consistency of these findings suggested that the premium reflects a fundamental phenomenon, which prompted theoretical research aimed at providing conceptual explanations for its existence.

Finance literature proposes a number of potential explanations for the size effect. Banz (1981) acknowledges that there are no theoretical foundations for the size effect he documents but he offers several conjectures consistent with the observed empirical pattern. Referring to the Klein and Bawa's (1977) model that shows that due to the estimation risk investors may be reluctant to invest in securities with limited available information, Banz (1981) proposes that the higher realized returns of small stocks may reflect a compensation for their higher information risk due to the lower quality of information disclosures provided by small firms.

Another commonly proposed explanation is based on the idea that small firms have a greater risk of financial distress (Chan and Chen 1991; Fama and French 1992, 1993, 1995, 1996; Chen and Zhang 1998; Vassalou and Xing 2004). The CAPM framework assumes away the cost of bankruptcy and financial distress. However, these costs can be significant especially for firms rich in intangible assets that cannot be easily transferred to new owners. Low market capitalization (i.e. small firm size) may result from a recent decrease in a firm's stock price that may be motivated by investors' concerns about the firm's financial viability. Hence, the higher

realized returns on small firms may compensate investors for the higher expected distress risk for firms in financial difficulty. Consistent with this proposition Vassalou and Xing (2004) show that the size effect is only significant in the quintile of stocks with the highest default risk and Hwang et al. (2010) find that the size effect can be explained by the excess credit spread on a firm's bonds that can be seen as a proxy for a firm's default risk. In contrast, Dichev (1998) finds that contrary to the expectations firms with high bankruptcy risk earn lower stock returns, which casts doubt on the distress risk explanation.

Furthermore, the size premium may represent a compensation for market imperfections that impair liquidity of small stocks to a greater degree than large stocks (Amihud and Mendelson 1986; Datar, Naik, and Radcliffe 1998; Amihud 2002; Pástor and Stambaugh 2003; Acharya and Pedersen 2005; Liu 2006). Liquidity of a stock is typically defined as its ability to be readily converted into cash at low cost (Liu 2006). Relative trading costs are likely larger for small stocks than for large stocks due to both the wider relative bid-ask spreads and the greater impact that a transaction of a given magnitude has on the stock price. The liquidity considerations may be particularly pressing for large institutional investors who trade large volumes of stocks and who may therefore be prone to avoid investment in small stocks. If large investors shy away from small stocks they constrain the demand, which ultimately depresses their price to the point where the higher expected return provides a fair compensation to the investors for the lower liquidity of small stocks.

The pervasiveness of the empirical evidence on the size effect together with the plausibility of the proposed theoretical explanations has gradually established firm size as a pricing factor featured in several empirical asset pricing models, e.g. the three-factor model (Fama and French 1993), the four-factor model (Carhart 1997), and the five-factor model (Fama and French 2015) that are a *de facto* industry standard for estimating expected stock returns.

Nevertheless, firm size was not *ex ante* envisaged as a pricing factor in asset pricing theory. Instead, the conceptual arguments of why it proxies for risk were provided only *ex post* when its predictive power was documented empirically. Consequently, there is still a controversy over whether the empirical findings should be interpreted as evidence on the ability of firm size to capture underlying risk characteristics. Horowitz *et al.* (2000a) argue that “despite the widespread use of size, the empirical basis for this practice is questionable” (p.84).

The size premium exhibits several puzzling characteristics that fuel this skepticism. First, empirical research suggests that the size premium significantly varies across regions and for some regions the predictive power of firm size is overshadowed by other pricing factor candidates. Hou *et al.* (2011) investigate stock returns in 49 economies, they reject the covariance risk model of firm size, and they 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 they conclude that both firm size and stock price momentum fail to reliably predict future stock returns. These findings are problematic because if firm size proxies for underlying risk, it should be systematically associated with stock returns across various contexts.

Second, some studies suggest that the size premium disappeared after the 1980s and then resurfaced after 2000 (Dichev 1998; Horowitz *et al.* 2000b; van Dijk 2011). This finding also constitutes a challenge to the interpretation of firm size as a risk proxy as these should be relatively stable over time. Researchers disagree over how these empirical findings should be interpreted. For example, Horowitz *et al.* (2000a) consider the disappearing size effect as evidence inconsistent with the proposition that firm size captures a hidden risk dimension. In contrast, Hou and van Dijk (2014) suggest that the disappearance of the size effect from *realized* stock returns does not necessarily imply a change in *expected* returns. They document that small

(large) firms experienced negative (positive) profitability shocks after the early 1980s that affected realized stock returns. However, after adjusting for the effect of these profitability shocks the difference in expected return on small and large stocks (i.e. the expected size premium) persists.

Third, the size premium is concentrated in the month of January (Keim 1983). Considering firm size as a risk proxy it is not trivial to explain why small stocks are systematically riskier in January while they are no more risky than other stocks in the remaining months of the year. 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 be contrary to the notion of systematic risk, which predicts low returns on riskier assets in times of economic downturn when scarcity of income increases.

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, the risk-based explanation of the firm size does not seem to be entirely compatible with another commonly proposed risk proxy the stock price “momentum” (Conrad and Kaul 1998; Chordia and Shivakumar 2002; Carhart 1997). Firm size is typically operationalized as a natural logarithm of a firm’s market capitalization. Momentum tends to be measured as the past 6-month dividend-adjusted stock return and so a positive stock price momentum implies an increase in firm size. It is challenging to reconcile the notion of small

stocks being riskier while at the same time stock that have recently become smaller (i.e. stocks with a negative stock price momentum) are less risky.

These puzzling findings challenge the role of firm size as a risk proxy and cast doubt on its use as a pricing factor. Despite of the *ex post* rationalizations of what risk dimensions firm size captures it is conceivable that the reported associations are spurious and they result from flawed methodologies, data mining, extreme observations, or other sources of bias. For example, Lo and MacKinlay (1990), MacKinlay (1995), and Berk (2000) argue that when stocks are sorted into portfolios based on previously empirically documented characteristics, conventional tests may overstate the statistical significance of the association between firm characteristics and stock returns. Due to the extensive use of firm size as a risk factor in the empirical asset pricing models it is important to investigate the impact these biases may have on the magnitude of the estimated coefficients. To that end we systematically analyze findings on the association between firm size and stock returns reported in prior literature and we provide new insights about the validity of firm size as a risk proxy and about the magnitude of the size premium.

3. Research Design

3.1. Methodology

In asset pricing, firm size is traditionally measured as a firm's market value of equity computed as the firm's stock price times the number of shares outstanding. The measure is bounded at zero and it is right-skewed and so it is common to log-transform it to get values of firm size that are closer to a normal distribution. The size effect has traditionally been estimated using two main approaches (Crain 2011; Fama and French 2008). The first approach uses the raw

firm size as the explanatory variable for the cross-sectional variation in stock returns (e.g. Fama and French 1992). Testing the relationship between firm size and stock returns is subject to several econometric challenges. While long-run variation in stock returns is plausibly driven by the differences in individual firms' fundamentals (i.e. a firm's systematic risk), in a short-run realized stock returns are rather volatile.

The noise caused by stock price oscillation around the intrinsic value entails large standard deviations of estimated coefficients, which makes it hard to empirically discern any relationship between stock returns and fundamentals (Merton 1980). In addition, some of the explanatory variables, namely the market beta of the CAPM, are not observable, they must be estimated and so their measurement is contaminated with an estimation error that may be correlated with the error term of the cross-sectional regression and so it may bias the estimated slope coefficients (Blume 1970).

To address these issues the cross-sectional regressions are typically run separately for every month and the time-series mean and variance in the estimated monthly slope coefficients are then used to compute the test statistics (Fama and MacBeth 1973, henceforth, the FM regressions). In addition, to amplify the signal-to-noise ratio some researchers pre-sort stocks into portfolios based on the explanatory variables and they use in the subsequent empirical tests the returns of the portfolios rather than of the individual stocks. Grouping stocks into portfolios washes away some of the idiosyncratic variation in individual stocks' returns, which increases the power of the empirical tests. At the same time, however, it reduces the number of observations and it may potentially eliminate some of the legitimate variation in stock returns driven by a firm's fundamentals.

The explanatory power of individual variables is tested in several steps. First, the unobservable variables, such as the market β , are estimated in time-series for the individual

firms using equation (1).

$$R_{it} = \alpha_i + \beta_i R_{mt} + \epsilon_{it}, \quad (1)$$

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.

Subsequently, the market beta estimated using equation (1) is used together with additional explanatory variables (including firm size) as right-hand-side variables in the monthly cross-section regressions following equation (2).

$$R_{it} = \gamma_{0t} + \gamma_{1t} \hat{\beta}_i + \gamma_{2t} X_i + u_{it}, \quad (2)$$

where $\hat{\beta}_i$ – beta estimated from (1),

X_i – vector of additional explanatory variables (including size).

The second approach assumes that firm characteristics such as firm size proxy for hidden risk dimensions and it suggests that stock returns are determined by a firm's exposure to the variation in risk premiums associated with these characteristics (Fama and French 1993). This approach requires a construction of mimicking portfolios representing the underlying pricing factors. The small-minus-big (*SMB*) factor that is based on firm size is defined as the difference in average monthly stock returns on a group (typically a quintile) of the smallest stocks less the equivalent stock returns on a group of the largest stocks. To capture individual firms' exposure to the *SMB* factor their monthly stock returns are regressed on the time-series of the *SMB* monthly premiums and the factor loadings of individual firms are approximated by the slope coefficients from these time-series regressions. The factor loadings are subsequently used as explanatory variables in monthly cross-sectional Fama-MacBeth (1973) regressions

that test whether the differences in individual firms' sensitivities to the pricing factor explain the cross-section of stock returns.

The second way to estimate the size effect, employed, e.g., in Easterday et al. (2009), Horowitz et al. (2000b), 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 follow the former approach which estimates the size effect by regressing stock returns on the natural logarithm of market value of equity or on other proxies for firm size.² While most studies use the cross-sectional FM regressions (a recent overview of such studies is available in Harvey et al. 2016) we do not restrict our sample to studies that employ this technique, but instead we include results from research articles reporting any form of a regression of stock returns on firm size.

We have several reasons for analyzing the studies that use firm size rather than the sensitivity to the *SMB* factor. First, the latter approach that estimates the sensitivity to the *SMB* factor implicitly assumes that used firm characteristics are valid proxies for hidden risk dimensions. In contrast, the former approach is more fundamental as it views the validity of firm characteristics as risk proxies as an open question. Second, most studies using the latter approach use mimicking portfolio returns published on Kenneth French's website, which makes the studies more homogeneous and thereby less suitable for a meta-analysis.

Third, the latter approach extensively relies on sorting stocks into portfolios and the

² 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.

estimated mean portfolio returns are sometimes reported without the corresponding t -statistics (or standard errors) that we require for a meta-analysis. Fourth, while the use of portfolios is understandable due to the noisy stock returns (see above) it comes at a cost of (i) eliminating a great deal of variation in the data much of which may be driven by fundamentals, and (ii) severely limiting the number of explanatory variables that constitute the dimensions of the sort. Thus sorting to portfolios has been extensively criticized (Lo and MacKinlay 1990; Berk 2000; Fama and French 2008; van Dijk 2011). Fama and French (2008) provide an argument against using sorts for making inferences about the marginal effects of individual stock return determinants:

... 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)

We collect point estimates and standard errors from the studies that satisfy the criteria described above to test for the selective reporting bias. Many authors recognize this bias as a serious issue in economics research (Ashenfelter and Greenstone 2004; Stanley 2005; Havranek 2015; Havranek and Kokes, 2015; Havranek et al. 2018a; Havranek et al. 2018b; Havranek et al. 2018c) that arises from researchers' tendency to primarily report the estimates that display the sign predicted by the theory and consistent with prior studies, along with those that are statistically significant.

Our empirical test of the publication selection bias exploits the association between the estimated slope coefficient and its standard error. Absent a publication bias, σ coefficient from equation (3) below is expected to be equal to zero because the ratio of the reported estimate to its standard error follows the t-distribution. In contrast, if researchers are prone to report statistically significant negative estimates and to discard positive or insignificant estimates, large negative point estimates are more likely to be reported together with large standard errors, which motivates the prediction for the coefficient σ in regression (3) will be negative. A test of the hypothesis $H_0: \sigma = 0$ is known as the funnel-asymmetry test because it can also be examined visually by evaluating the symmetry of the corresponding scatter plot (Stanley 2005; Feld and Heckemeyer 2011; Feld et al. 2013; see also the useful guidelines by Stanley et al., 2013). Coefficient S_0 from equation (3) represents the underlying mean effect corrected for the reporting bias:

$$S_{it} = S_0 + \sigma \cdot SE(S_{it}) + \epsilon_{it}, \quad (3)$$

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

We estimate equation (3) using several approaches. First, we use the ordinary least squares (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 (3) 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 of the equation (4), with weights based on variable $1/SE(S_{it})$, which we call *precision*. In addition, following Irsova and Havranek (2013) and Havranek et al. (2015), we run specification (4) 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:

$$\frac{S_{it}}{SE(S_{it})} = S_0 \frac{1}{SE(S_{it})} + \sigma + u_{it} \quad (4)$$

In equation (4), 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 -1.96, the most commonly used threshold for statistical significance. 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.

We further acknowledge that some omitted study characteristics, such as the use of various estimation techniques, can affect both the estimates of slope coefficients and their standard errors. To address this endogeneity concern, we estimate the equation (4) using the instrumental variable regression technique. We argue that the number of observations n used for estimating a particular coefficient is by definition correlated with the reported standard error but it is plausibly independent of the choice of the estimation technique. We use the reciprocal of the square root of the number of observations $1/\sqrt{n}$ (Stanley 2005; Zigrainova and Havranek 2016; Havranek and Irsova 2017; Hampl and Havranek, 2019) as an instrument in estimating the equation (4).

Next, we partition our sample using measures that we expect to affect the publication selection bias. We include in specification (3) interaction terms of the standard error with the recursive impact factor of the journal where the article containing the estimate is published (as reported on the IDEAS/RePEc website³) and with the year of its publication. The direction of

³ 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,

the effect of journal quality on the strength of publication selection bias is not obvious. On the one hand, higher-quality journals employ more stringent review procedures, decreasing the likelihood of publishing studies with heavy data mining and back-fitting of the results. On the other hand, researchers submitting studies to high-quality journals may feel obliged to omit inconclusive or contra-intuitive findings that would render the empirical support of their paper less convincing and compromise the odds of being accepted for publication. Similarly, it is not clear how the year of publication should affect the selective reporting bias. On the one hand, recent studies employ more sophisticated econometric techniques, building on advances in the state-of-art methodology, which should bring to the estimated effect closer to the true effect. On the other hand, more recent research may have a greater tendency to report results consistent with the larger pool of already published findings.

As an additional test of selective reporting, we employ Hedges' (1992) model for detecting the 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 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 (which in our case represents the publication of a study) is a function of p-value, and the abovementioned

see, for example, Havranek and Irsova (2011).

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*

To collect our sample of the published estimates we first perform a Google Scholar search using keywords *(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.* We consider the first 500 results provided by Google Scholar sorted by relevance. We only consider studies published in peer-reviewed journals that have passed the quality control mechanism of the review process. Rusnak et al. (2013) find no difference in the extent of selective reporting between published and unpublished studies in economics and they argue that because the authors' ultimate aim is to publish their papers in research journals the incentives discard non-intuitive findings applies equally to published and non-published research. (A similar approach is employed by Havranek and Rusnak, 2013, and Havranek and Sokolova, 2019.) We then "snowball" our sample by considering all the references (i) from the ten most-cited papers collected via Google Scholar search, and (ii) from the prominent review papers (Crain 2011; van Dijk 2011; Harvey et al. 2016; Horowitz et al. 2000a; Hou et al. 2011).

We start our sample period in 1981, which is the year of Banz's (1981) initial study on size effect, and terminate our search on March 31, 2017. 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. Even though most of the studies we source our estimates from use the FM approach we do not restrict our dataset to this specific methodology. We collect all estimates reported in the studies that meet our search criteria. 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. Our data and code are available at meta-analysis.cz/size.

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			

We report summary statistics of our dataset in Table 1. The mean reported estimate of coefficient $\hat{\gamma}_{2t}$ from regression (2) is -0.092, which is consistent with the conventional notion that small stocks tend to outperform large stocks.

Table 2 shows basic data partitions. The estimates that we analyze exhibit 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 time period and the region. In the table we only show estimates that use data solely from the time period mentioned on the left: for example, there are 16 estimates that only use data after 2000 and no older data. 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. (2000a). In addition, the size effect is somewhat less pronounced in the USA than in other regions used in the data sample. Consistent with Keim (1983) the size effect is concentrated in the month of January. 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
<i>Year of observations</i>					
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
<i>Geographical region</i>					
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
<i>Stock exchange (US)</i>					
NYSE only	305	-0.031	0.123	-1.19	0.21
Any	784	-0.100	0.558	-3.60	4.45
<i>Month of returns</i>					
January	113	-0.481	0.851	-5.94	0.92
Non-January	1633	-0.043	0.236	-1.80	0.52
<i>Estimation technique</i>					
OLS	928	-0.114	0.543	-5.94	4.69
Other	818	-0.068	0.400	-3.60	1.48
<i>Stock returns</i>					
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, Hampl et al. 2019). 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 (Dissanaïke 2002).

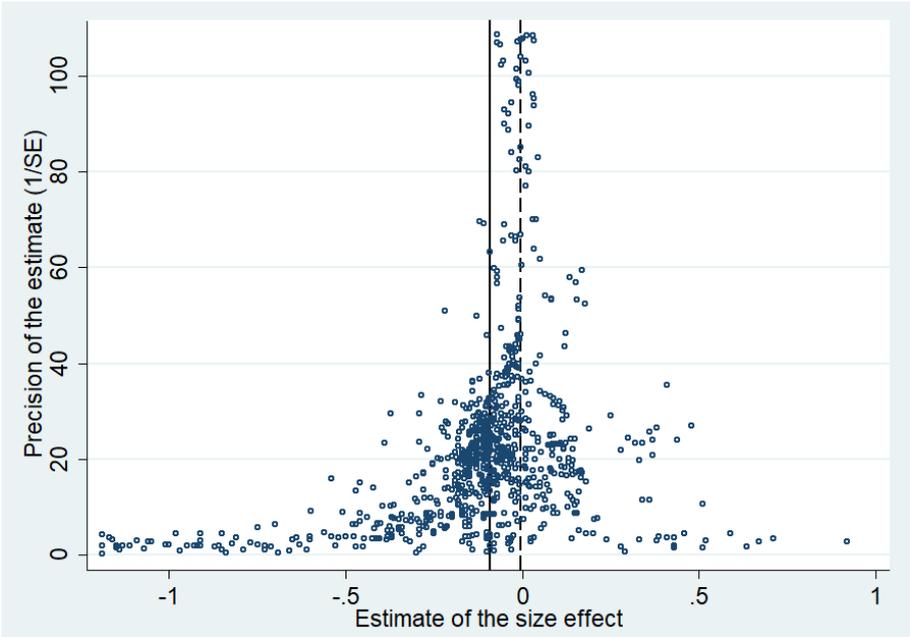
2. *Different measures of variability.* Different authors use different adjustments of standard errors and t-statistic to account for autocorrelation (Asparouhova et al. 2013), measurement errors (Fletcher 1997), and heteroscedasticity (Chen et al. 2002), etc.
3. *Insufficient reporting.* During the 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

The funnel plot shown in Figure 1 visualizes the effect of potential selective reporting using the method 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 a 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 would be dispersed around the mean. Using our data on size premium, we find the plot to be asymmetric and the individual point estimates to be concentrated in the left-hand tail. Therefore, it seems that positive estimates are less likely to be reported in published academic studies and if they are reported their precision tends to be fairly low (or equivalently, their standard errors tend to be large). The pattern in Figure 1 is thus consistent with selectivity in reporting the slope coefficient estimates.

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.

To formally test for publication bias we run regressions following the specifications described in Section 3. The results from this estimation are reported in Table 3. Column (1) presents the baseline OLS result of regressing the estimated slope coefficients at firm size on their standard errors. The significantly negative σ coefficient indicates the presence of the selective reporting bias. The estimated intercept of -0.032 represents the underlying mean size effect corrected for the selective reporting bias. Also consistent with selective reporting, the value of the intercept is much smaller than the un-adjusted mean slope coefficient reported in descriptive statistics in Table 1. The baseline result thus provides support for the existence of the negative size effect but at the same time suggests that the effect is less pronounced than commonly argued.

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*** (0.0862)	-0.897*** (0.127)	-0.526*** (0.102)	-1.159*** (0.0300)	-0.554*** (0.165)	-0.494 (0.311)
Constant	-0.0315*** (0.00814)	-0.0236** (0.0113)	-0.0387 (0.0245)	-0.000291 (0.000271)	-0.0358*** (0.0128)	-0.0611** (0.0277)
Observations	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 regressions with fixed effects and with between effects respectively. Both specifications indicate the presence of a selective reporting bias. The bias seems to be stronger when within-study variation is considered. This finding is particularly salient when within-study variation is considered, which is consistent

with strategic reporting of significantly negative coefficients. Column (4) of Table 3 reports the results estimating specification (4) using WLS when the precision variable is used to weight the observations. WLS produces the most pronounced publication bias. At the same time, the mean slope coefficient corrected for this bias remains negative but it ceases to be significant. Similar, albeit weaker 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. Overall, these results provide evidence consistent with selective reporting bias in the estimates of the size effect.

The last column (6) 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 from the instrumental variable regression yield signs consistent with our earlier results. Nevertheless, as the IV estimation is less precise the coefficient of SE is statistically insignificant.

Next, to investigate the pattern of publication bias, we follow Havranek and Irsova (2012), among others, and interact the reported standard error with the recursive impact factor of the journal where the study is published and with the year of its publication. We present the results based on the OLS and the FE estimation Table 4.⁴ 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

⁴ The results of the other estimation approaches are shown in Appendix B.

sign in the fixed effects regression; in both specifications, the coefficient is either insignificant or (in one case) only mildly significant. 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*** (0.175)	-1.294*** (0.156)	-1.288*** (0.128)	-1.096*** (0.169)	-1.272*** (0.178)	-1.433*** (0.202)
SE*Impact	-0.0686 (0.0603)		-0.185** (0.0766)	0.247 (0.170)		0.234 (0.159)
SE*Pub. Year		0.000164*** (0.0000274)	0.000239*** (0.0000460)		0.000135*** (0.0000393)	0.000124*** (0.0000449)
Constant	-0.0351*** (0.00771)	-0.0241*** (0.00857)	-0.0247*** (0.00647)	-0.0305** (0.0153)	-0.0197* (0.0108)	-0.0266* (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 always positive and highly significant, suggesting that selective reporting is less pronounced in more recent studies. This results may be due to the use of more refined and precise econometric techniques or due to a growing acceptance of positive or nonsignificant estimates following the alleged “disappearance” of the size effect in the US data after the evidence on was first published in the 1980s (Horowitz et al. (2000a).

Finally, 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, the model tests whether publication probability depends on the reported level of statistical

significance. The results of the estimation are reported in Table 5, where weight ω_1 , associated with a p-value of 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. Typically the annual size premium is defined as the difference in stock returns between the 10th and 1st deciles when the portfolio breakpoints are set according to the NYSE stocks market capitalization. The universe of U.S. stocks is typically sourced from the CRSP database and it comprises all stocks

traded on the NYSE, NYSE-MKT (formerly AMEX) and NASDAQ, excluding the closed-end funds, preferred stocks, REITs, foreign stocks and trusts. The size premium is computed as the difference between mean annualized monthly excess stocks returns (over the CAPM) on the portfolio of the smallest and the largest stocks.

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 90 th percentile (size premium), unadjusted	Annualized difference with 90 th percentile (size premium), adjusted for selective reporting bias
5 th	5.14	5.88%	1.98%
10 th	5.83	5.08%	1.72%
15 th	6.23	4.62%	1.56%
20 th	6.61	4.18%	1.42%
25 th	6.87	3.88%	1.31%
30 th	7.13	3.59%	1.22%
35 th	7.37	3.31%	1.12%
40 th	7.57	3.09%	1.05%
45 th	7.79	2.84%	0.97%
50 th	8.01	2.59%	0.88%
55 th	8.19	2.38%	0.81%
60 th	8.38	2.18%	0.74%
65 th	8.64	1.88%	0.64%
70 th	8.90	1.58%	0.54%
75 th	9.19	1.26%	0.43%
80 th	9.50	0.91%	0.31%
85 th	9.88	0.50%	0.17%
90 th	10.33	0.00%	0.00%
95 th	10.93	-0.66%	-0.23%
100 th	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 6.⁵ The size premium can be calculated as the

⁵ Available at http://mba.tuck.dartmouth.edu/pages/faculty/ken.french/data_library.html. The latest

difference between the slope coefficients adjusted for selective reporting bias 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.

As a benchmark for the un-adjusted size premium we use the simple mean reported coefficient of -0.092 reported in Table 1 and compute the difference between the implied return on the 10th and the 90th percentile of NYSE stocks. We obtain a benchmark monthly size premium of 0.415%, or 5.08% annualized, which is broadly consistent with the magnitude of the historical size premium that tends to be proposed in prior literature. To calculate the size premium, we use the estimate of -0.032 from our baseline specification (1) reported in Table 3, which represents our estimate of association between firm size and stock returns corrected for the effect of selective reporting (note that the size premium would be much smaller if we selected the WLS result). 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.

To assess the robustness of our results to changes in methodology we repeat these computations excluding observations of the size effect with the returns being non-monthly, focusing on US stocks only. This methodological modification reduces the number of sampled coefficients to 946. In this restricted sample the mean slope coefficient is -0.095. The adjusted value of size effect from specification (1) in Table 3 is -0.0259, which implies an unadjusted annualized size premium of 5.25%. In comparison, adjusting for the publication selection bias reduces the estimated magnitude of the size premium to 1.41%. Next, we re-estimate our main specification for the subsample of 876 slope coefficients based on regressions that control for

available data were as of May 2017.

the market beta and for the remaining 870 slope coefficients that do not control for market beta. We observe a slightly weaker size effect for coefficients from regressions that control for the market beta, but the tendency towards selective reporting is strongly statistically significant (at 1 percent level) in both subsamples (not tabulated). Hence, consistent with our previous results this result supports the smaller magnitude of the size premium after adjusting for the selective reporting bias.

Finally, in Appendix C we also recompute our main tables (Table 3 and Table 4) after collecting data from working papers. As we have noted, in the main body of the paper we focus on published studies, because we expect peer review to improve the quality of presented analyses and because published studies are less likely to contain errors that could increase the variance of our meta-analysis estimators. To search for working papers, we use the same query as for the main data set but restrict our attention to working papers presented in 2010 or later with at least 10 citations in Google Scholar. In the end we include 10 new studies, which provide 167 estimates of the size effect, all of them using data extending after year 2000. The tables in Appendix C show that our conclusions regarding the magnitude of publication bias and the mean size effect (and thus also the corresponding size premium) remain unchanged after including the working papers.

5. Conclusion

We perform a meta-analysis of 1,746 estimates from 102 studies analyzing the impact of firm size on stock returns. We use three approaches 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 approaches 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. Assuming a linear association between firm size and its stock returns our estimate of the difference in annual stock returns on the smallest and largest market capitalization quintile based on the NYSE cutoffs is 1.72% after correcting for this bias. 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. We observe a drop in selective reporting 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, which implies less need to discard nonsignificant or even positive estimates of the size effect, thereby alleviating publication bias.

Our study is subject to several caveats. First, our conclusions are based on the assumption that *ex post* realized stock returns are on average an unbiased proxy for returns expected *ex ante* (see Hou and van Dijk 2014 for a discussion). Assuming that investors are rational in their expectations realized returns should converge to expected returns in a long run. However, over shorter time horizons realized returns may significantly deviate from the expectations. The proposition that the size premium compensates investors for higher risk of small stocks implies that even though on average small stocks should outperform large stocks, returns on small stocks may be inferior to large stocks over several years (which is, after all, the reason why investment in small stocks is considered risky). The use of realized returns as a proxy for expected returns is thus problematic especially in studies that use short sample periods that may be prone to incorrect inferences due to unrepresentative samples.

Second, our findings support the proposition that the magnitude of the size premium varies over time (Horowitz et al. 2000a; van Dijk 2011) and more specifically that it has decreased after 1980s. Nevertheless, we acknowledge that our analysis is constrained by the sample periods covered in the articles that constitute our sample. If the size premium has recently re-surged as suggested for example by van Dijk (2011) our meta-analysis may not be able to detect this shift because of the limited number of published articles that cover these recent years. Future research should investigate how the size premium varies over time and how it depends on time-varying measures such as the phases of economic cycle.

Third, our approach of estimating the magnitude of the size premium is based on NYSE market capitalization cutoffs and it assumes a linear association between firm size and stock returns. Prior research suggests that the premium is concentrated in the sub-sample of very small stocks. Hence, our size premium estimate is applicable for firms of sizes comparable with the NYSE market capitalization quintiles. Our findings are not necessarily generalizable to

micro stocks at the bottom or below the smallest NYSE quintile.

Fourth, the preponderance of studies in our sample are based on the U.S. stock market data. These markets are characterized by strong institutional regulation, effective corporate governance enforcement and high stock liquidity (La Porta, Lopez-de-Silanes, and Shleifer 1999; La Porta et al. 2000, 2002). If the size premium reflects a compensation for a greater information risk, lower liquidity and/or higher sensitivity of small stock returns to market imperfections our findings may not be directly generalizable to other less well-organized markets. In market with weaker institutional framework we expect the size premium to be larger. Further research is needed on the variation of the size premium across different geographic markets and on its dependence on the quality of the institutional setting.

Fifth, the size effect reported in the studies we analyze may be affected by the inclusion of control variables and by the estimation procedure. For example, we observe a slightly weaker size effect and a stronger tendency towards selective reporting in estimates that control for the market beta. We are currently working on a paper that would investigate what impact variation in methodology used to estimate the sampled coefficients has on the strength of the documented association and on the tendency to report results selectively.

References

- Acharya, Viral V. and Lasse Heje Pedersen. 2005. "Asset Pricing with Liquidity Risk." *Journal of Financial Economics* 77 (2), 375.
- Amihud, Yakov. 2002. "Illiquidity and Stock Returns: Cross-Section and Time-Series Effects." *Journal of Financial Markets* 5 (1), 31–56.
- Amihud, Yakov and Haim Mendelson. 1986. "Asset Pricing and the Bid-Ask Spread." *Journal of Financial Economics* 17 (2), 223–49.
- 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.
- 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), 407-427.
- 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.
- Blume, Marshall E. 1970. “Portfolio Theory: A Step Toward Its Practical Application.” *The Journal of Business* 43 (2), 152–73.
- 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, Nai-fu and Feng Zhang. 1998. “Risk and Return of Value Stocks.” *The Journal of Business* 71 (4), 501–35.
- 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), 985-1019.
- Conrad, Jennifer and Gautam Kaul. 1998. “An Anatomy of Trading Strategies.” *Review of Financial Studies* 11 (3), 489-519.
- Crain, Michael A. 2011. *A Literature Review of the Size Effect*. Rochester, NY: Social Science Research Network.
- 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.
- 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.
- Doucouliafos, Chris and Tom D. Stanley. 2013. “Are All Economic Facts Greatly Exaggerated? Theory Competition and Selectivity.” *Journal of Economic Surveys* 27 (2), 316–339.
- 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. 1995. "Size and Book-to-Market Factors in Earnings and Returns." *Journal of Finance* 50 (1), 131-155.
- Fama, Eugene F. and Kenneth R. French. 1996. "Multifactor Explanations of Asset Pricing Anomalies." *Journal of Finance* 51 (1), 55–84.
- 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." *Journal of Political Economy* 81 (3), 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.
- Hampl, Mojmír and Tomas Havranek. 2019. "Central Bank Equity as an Instrument of Monetary Policy." *Comparative Economic Studies*, forthcoming.
- Hampl, Mojmír, Tomas Havranek, and Zuzana Irsova. 2019. "Foreign Capital and Domestic Productivity in the Czech Republic: A Meta-Regression Analysis." *Applied Economics*, forthcoming.
- 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, Dominik Herman, and Zuzana Irsova. 2018a. "Does Daylight Saving Save Electricity? A Meta-Analysis." *Energy Journal* 39(2), 35-61.
- 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, Zuzana Irsova, and Tomas Vlach. 2018b. "Measuring the Income Elasticity of Water Demand: The Importance of Publication and Endogeneity Biases." *Land Economics* 94(2), 259-283.
- Havranek, Tomas, Zuzana Irsova, and Olesia Zeynalova. 2018c. "Tuition Fees and University Enrollment: A

- Meta-Regression Analysis.” *Oxford Bulletin of Economics and Statistics* 80(6), 1145-1184.
- Havranek, Tomas and Ondrej Kokes. 2015. “Income Elasticity of Gasoline Demand: A Meta-Analysis.” *Energy Economics* 47(1), 77-86.
- Havranek, Tomas and Marek Rusnak. 2013. “Transmission Lags of Monetary Policy: A Meta-Analysis.” *International Journal of Central Banking* 9(4), 39-75.
- Havranek, Tomas, Marek Rusnak, and Anna Sokolova. 2017. "Habit Formation in Consumption: A Meta-Analysis." *European Economic Review* 95, 142-167.
- Havranek, Tomas and Anna Sokolova. 2019. “Do Consumers Really Follow a Rule of Thumb? Three Thousand Estimates from 144 Studies Say "Probably Not".” *Review of Economic Dynamics*, forthcoming.
- Hedges, Larry V. 1992. “Modeling Publication Selection Effects in Meta-Analysis.” *Statistical Science* 246–255.
- Horowitz, Joel L., Tim Loughran, and N. E. Savin. 2000a. “The Disappearing Size Effect.” *Research in Economics* 54 (1), 83–100.
- Horowitz, Joel L, Tim Loughran, and N. E. Savin. 2000b. “Three Analyses of the Firm Size Premium.” *Journal of Empirical Finance* 7 (2), 143–53.
- Hou, Kewei and Mathijs A. van Dijk. 2014. *Resurrecting the Size Effect: Firm Size, Profitability Shocks, and Expected Stock Returns*. Rochester, NY: Social Science Research Network.
- 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.
- Hwang, Young-Soon, Hong-Ghi Min, Judith A. McDonald, Hwagyun Kim, and Bong-Han Kim. 2010. “Using the Credit Spread as an Option-Risk Factor: Size and Value Effects in CAPM.” *Journal of Banking & Finance* 34 (12), 2995–3009.
- Irsova, Zuzana and Tomas Havranek. 2013. “Determinants of Horizontal Spillovers from FDI: Evidence from a Large Meta-Analysis.” *World Development* 42(1), 1–15.
- 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.
- Jegadeesh, Narasimhan and Sheridan Titman. 2001. “Profitability of Momentum Strategies: An Evaluation of Alternative Explanations.” *Journal of Finance* Apr2001, Vol. 56 Issue 2.
- Keim, Donald B. 1983. “Size-Related Anomalies and Stock Return Seasonality.” *Journal of Financial Economics* 12 (1), 13–32.
- Klein, Roger W. and Vijay S. Bawa. 1977. “The Effect of Limited Information and Estimation Risk on Optimal Portfolio Diversification.” *Journal of Financial Economics* 5 (1), 89–111.
- 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.
- La Porta, Rafael, Florencio Lopez-de-Silanes, and Andrei Shleifer. 1999. “Corporate Ownership around the World.” *The journal of finance* 54 (2), 471–517.
- La Porta, Rafael, Florencio Lopez-de-Silanes, Andrei Shleifer, and Robert Vishny. 2000. “Investor Protection and Corporate Governance.” *Journal of Financial Economics* 58 (1), 3–27.
- La Porta, Rafael, Florencio Lopez-De-Silanes, Andrei Shleifer, and Robert Vishny. 2002. “Investor Protection and

- Corporate Valuation.” *The Journal of Finance* 57 (3), 1147–70.
- Lakonishok, Josef, Andrei Shleifer, and Robert W. Vishny. 1994. “Contrarian Investment, Extrapolation, and Risk.” *The Journal of Finance* 49 (5), 1541–1578.
- 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), 13-37.
- Liu, Weimin. 2006. “A Liquidity-Augmented Capital Asset Pricing Model.” *Journal of Financial Economics* 82 (3), 631–71.
- 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.
- Merton, Robert C. 1980. “On Estimating the Expected Return on the Market: An Exploratory Investigation.” *Journal of Financial Economics* 8 (4), 323–61.
- Pástor, Luboš and Robert F. Stambaugh. 2003. “Liquidity Risk and Expected Stock Returns.” *Journal of Political Economy* 111 (3), 642–85.
- 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), 425-442.
- 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.
- Stanley, Tom D., Hristos Doucouliagos, Margaret Giles, Jost H. Heckemeyer, Robert J. Johnston, Patrice Laroche, Jon P. Nelson, Martin Paldam, Jacques Poot, Geoff Pugh, Randall S. Rosenberger, and Katja Rost. 2013. “Meta-Analysis of Economics Research Reporting Guidelines.” *Journal of Economic Surveys* 27(2), 390-394.
- 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.” <i>Journal of financial Economics</i> 77 (2), 375–410.	2676	1.498
Amel-Zadeh, Amir. 2011. “The Return of the Size Anomaly: Evidence from the German Stock Market.” <i>European Financial Management</i> 17 (1), 145–182.	30	0.162
Amihud, Yakov. 2002. “Illiquidity and Stock Returns: Cross-Section and Time-Series Effects.” <i>Journal of Financial Markets</i> 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.” <i>The Journal of Finance</i> 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.” <i>The Journal of Finance</i> 61 (1), 171–94.	286	1.776
Ang, Andrew, Joseph Chen, and Yuhang Xing. 2006. “Downside Risk.” <i>Review of Financial Studies</i> 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.” <i>Journal of Financial Economics</i> 91 (1), 1–23.	637	2.081
Asparouhova, Elena, Hendrik Bessembinder, and Ivalina Kalcheva. 2013. “Noisy Prices and Inference Regarding Returns.” <i>The Journal of Finance</i> 68 (2), 665–714.	77	1.776
Avramov, Doron and Tarun Chordia. 2006. “Asset Pricing Models and Financial Market Anomalies.” <i>Review of Financial Studies</i> 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.” <i>Journal of banking & finance</i> 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.” <i>Journal of Financial Economics</i> 99 (2), 427–446.	339	1.498
Banz, Rolf W. 1981. “The Relationship between Return and Market Value of Common Stocks.” <i>Journal of financial economics</i> 9 (1), 3–18.	5866	1.498
Barry, Christopher B. and Stephen J. Brown. 1984. “Differential Information and the Small Firm Effect.” <i>Journal of Financial Economics</i> 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.” <i>Emerging Markets Review</i> 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.” <i>European Financial Management</i> 16 (2), 165–190.	52	0.162
Bhandari, Laxmi Chand. 1988. “Debt/Equity Ratio and Expected Common Stock Returns: Empirical Evidence.” <i>The journal of finance</i> 43 (2), 507–528.	1308	1.776
Borys, Magdalena Morgese and Petr Zemčik. 2011. “Size and Value Effects in the Visegrad Countries.” <i>Emerging Markets Finance and Trade</i> 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." <i>Financial management</i> 35 (3), 97–116.	552	0.236
Brennan, Michael, Tarun Chordia, and Avanidhar Subrahmanyam. 2004. "Cross-Sectional Determinants of Expected Returns." <i>The Legacy of Fischer Black</i> 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." <i>Journal of Financial Economics</i> 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." <i>Journal of Financial Economics</i> 105 (3), 523–541.	37	n/a
Bryant, Paul S. and Venkat R. Eleswarapu. 1997. "Cross-Sectional Determinants of New Zealand Share Market Returns." <i>Accounting & Finance</i> 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." <i>Journal of financial economics</i> 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." <i>Journal of Business Finance & Accounting</i> 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." <i>The Journal of Finance</i> 43 (2), 309–325.	409	1.776
Chan, K. C. and Nai-Fu Chen. 1991. "Structural and Return Characteristics of Small and Large Firms." <i>The Journal of Finance</i> 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." <i>Journal of Financial Economics</i> 14 (3), 451–471.	180	2.081
Chan, Louis K. C., Yasushi Hamao, and Josef Lakonishok. 1991. "Fundamentals and Stock Returns in Japan." <i>The Journal of Finance</i> 46 (5), 1739–64.	90	1.776
Chan, Louis K. C., Yasushi Hamao, and Josef Lakonishok. 1993. "Can Fundamentals Predict Japanese Stock Returns?" <i>Financial Analysts Journal</i> 49 (4), 63–69.	1187	1.776
Chan, Louis K. C., Jason Karceski, and Josef Lakonishok. 1998. "The Risk and Return from Factors." <i>Journal of Financial and Quantitative Analysis</i> .	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." <i>Managerial Finance</i> 33 (8), 574–594.	1466	1.776
Chen, Joseph, Harrison Hong, and Jeremy C. Stein. 2002. "Breadth of Ownership and Stock Returns." <i>Journal of Financial Economics</i> 66 (2–3), 171–205.	903	1.776
Chopra, Navin, Josef Lakonishok, and Jay R. Ritter. 1992. "Measuring Abnormal Performance." <i>Journal of Financial Economics</i> 31 (2), 235–68.	48	0.044
Chordia, Tarun, Avanidhar Subrahmanyam, and V.Ravi Anshuman. 2001. "Trading Activity and Expected Stock Returns." <i>Journal of Financial Economics</i> 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." <i>Pacific-Basin Finance Journal</i> 6 (3–4), 275–93.	1828	1.776
Claessens, Stijn, Susmita Dasgupta, and Jack Glen. 1995. <i>The Cross-Section of Stock Returns: Evidence from the Emerging Markets</i> . 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." <i>The Journal of Finance</i> 63 (4), 1609–51.	256	1.776
Cooper, Michael J., Huseyin Gulen, and Alexei V. Ovtchinnikov. 2010. "Corporate Political Contributions and Stock Returns." <i>The Journal of Finance</i> 65 (2), 687–724.	45	0.096
Cremers, KJ Martijn, Vinay B. Nair, and Kose John. 2009. "Takeovers and the Cross-Section of Returns." <i>Review of Financial Studies</i> 22 (4), 1409–1445.	59	0.032
Da, Zhi. 2009. "Cash Flow, Consumption Risk, and the Cross-Section of Stock Returns." <i>The Journal of Finance</i> 64 (2), 923–956.	1102	1.776
Datar, Vinay T., Narayan Y. Naik, and Robert Radcliffe. 1998. "Liquidity and Stock Returns: An Alternative Test." <i>Journal of Financial Markets</i> 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." <i>Journal of Futures Markets</i> 28 (11), 1013–1039.	5	0.096
Dichev, Ilia D. 1998. "Is the Risk of Bankruptcy a Systematic Risk?" <i>The Journal of Finance</i> 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." <i>The Journal of Finance</i> 57 (5), 2113–41.	132	0.082
Dissanaike, Gishan. 2002. "Does the Size Effect Explain the UK Winner-Loser Effect?" <i>Journal of Business Finance & Accounting</i> 29 (1–2), 139–154.	702	1.498
Doeswijk, Ronald Q. 1997. "Contrarian Investment in the Dutch Stock Market." <i>De Economist</i> 145 (4), 573–598.	40	n/a
Easley, David, Soeren Hvidkjaer, and Maureen O'Hara. 2002. "Is Information Risk a Determinant of Asset Returns?" <i>The Journal of Finance</i> 57 (5), 2185–2221.	61	0.014
Eleswarapu, Venkat R. 1997. "Cost of Transacting and Expected Returns in the Nasdaq Market." <i>The Journal of Finance</i> 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." <i>Journal of Financial Economics</i> 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." <i>Journal of Financial Research</i> 21 (3), 277–91.	150	0.162
Estrada, Javier and Ana Paula Serra. 2005. "Risk and Return in Emerging Markets: Family Matters." <i>Journal of Multinational Financial Management</i> 15 (3), 257–272.	24	0.067
Fama, Eugene F. and Kenneth R. French. 1992. "The Cross-Section of Expected Stock Returns." <i>The Journal of Finance</i> 47 (2), 427–465.	92	0.082
Fama, Eugene F. and Kenneth R. French. 2008. "Dissecting Anomalies." <i>The Journal of Finance</i> 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 Empirical Exploration of Berk's Critique." <i>Journal of Financial Research</i> 28 (4), 503–518.	126	0.288
Ferson, Wayne E. and Campbell R. Harvey. 1999. "Conditioning Variables and the Cross Section of Stock Returns." <i>The Journal of Finance</i> 54 (4), 1325–60.	473	2.081
Fletcher, Jonathan. 1997. "An Examination of the Cross-Sectional Relationship of Beta and Return: UK Evidence." <i>Journal of Economics and Business</i> 49 (3), 211–221.	260	2.081
Fu, Fangjian. 2009. "Idiosyncratic Risk and the Cross-Section of Expected Stock Returns." <i>Journal of Financial Economics</i> 91 (1), 24–37.	174	0.381
Garza-Gómez, Xavier, Jiro Hodoshima, and Michio Kunimura. 1998. "Does Size Really Matter in Japan?" <i>Financial Analysts Journal</i> 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." <i>Accounting & Finance</i> 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." <i>Journal of Financial Economics</i> 96 (1), 56–79.	335	1.776
Herrera, Martin J. and Larry J. Lockwood. 1994. "The Size Effect in the Mexican Stock Market." <i>Journal of Banking & Finance</i> 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." <i>European Financial Management</i> 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." <i>Journal of International Financial Markets, Institutions and Money</i> 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." <i>Journal of Economics and Business</i> 52 (6), 515–33.	1732	1.776
Horowitz, Joel L., Tim Loughran, and N. E. Savin. 2000. "The Disappearing Size Effect." <i>Research in Economics</i> 54 (1), 83–100.	630	1.498
Horowitz, Joel L., Tim Loughran, and N. E. Savin. 2000. "Three Analyses of the Firm Size Premium." <i>Journal of Empirical Finance</i> 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." <i>Review of Financial Studies</i> 18 (3), 981–1020.	32	0.008
Hou, Kewei, G. Andrew Karolyi, and Bong-Chan Kho. 2011. "What Factors Drive Global Stock Returns?" <i>Review of Financial Studies</i> 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." <i>Journal of Accounting and Economics</i> 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." <i>Financial Review</i> 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." <i>Journal of Empirical Finance</i> 28, 139–50.	265	0.057
Jaffe, Jeffrey, Donald B. Keim, and Randolph Westerfield. 1989. "Earnings Yields, Market Values, and Stock Returns." <i>The Journal of Finance</i> 44 (1), 135–148.	561	1.776
Jagannathan, Ravi and Yong Wang. 2007. "Lazy Investors, Discretionary Consumption, and the Cross-Section of Stock Returns." <i>The Journal of Finance</i> 62 (4), 1623–1661.	207	1.776
Jagannathan, Ravi and Zhenyu Wang. 1996. "The Conditional CAPM and the Cross-Section of Expected Returns." <i>The Journal of Finance</i> 51 (1), 3–53.	2424	1.776
Jegadeesh, Narasimhan. 1992. "Does Market Risk Really Explain the Size Effect?" <i>The Journal of Financial and Quantitative Analysis</i> 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." <i>Journal of Financial Research</i> 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." <i>Journal of Financial and Quantitative Analysis</i> 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." <i>The Journal of Finance</i> 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." <i>The Journal of Finance</i> 50 (1), 185–224.	1279	1.776
La Porta, Rafael. 1996. "Expectations and the Cross-Section of Stock Returns." <i>The Journal of Finance</i> 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." <i>Journal of Banking & Finance</i> 10 (1), 115–132.	342	0.208
Lakonishok, Josef, Andrei Shleifer, and Robert W. Vishny. 1994. "Contrarian Investment, Extrapolation, and Risk." <i>The Journal of Finance</i> 49 (5), 1541–78.	4840	1.776
Loughran, Tim. 1997. "Book-to-Market across Firm Size, Exchange, and Seasonality: Is There an Effect?" <i>Journal of financial and quantitative analysis</i> 32 (03), 249–268.	3964	1.776
Loughran, Tim and Jay R. Ritter. 1995. "The New Issues Puzzle." <i>The Journal of Finance</i> 50 (1), 23–51.	401	0.38
Moskowitz, Tobias J. and Mark Grinblatt. 1999. "Do Industries Explain Momentum?" <i>The Journal of Finance</i> 54 (4), 1249–90.	1418	1.776
Nagel, Stefan. 2005. "Short Sales, Institutional Investors and the Cross-Section of Stock Returns." <i>Journal of Financial Economics</i> 78 (2), 277–309.	678	1.498
Novy-Marx, Robert. 2013. "The Other Side of Value: The Gross Profitability Premium." <i>Journal of Financial Economics</i> 108 (1), 1–28.	351	1.498
Ozgun Demirtas, K. and A. Burak Guner. 2008. "Can Overreaction Explain Part of the Size Premium?" <i>International Journal of Revenue Management</i> 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." <i>Journal of Accounting Research</i> 45 (2), 427–67.	209	0.304
Pettengill, Glenn, Sridhar Sundaram, and Ike Mathur. 2002. "Payment For Risk: Constant Beta Vs. Dual-Beta Models." <i>Financial Review</i> 37 (2), 123–135.	55	0.046
Phalippou, Ludovic. 2007. "Can Risk-Based Theories Explain the Value Premium?" <i>Review of Finance</i> 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." <i>Financial Services Review</i> 10 (1–4), 291–302.	23	0.034
Pontiff, Jeffrey and Artemiza Woodgate. 2008. "Share Issuance and Cross-Sectional Returns." <i>The Journal of Finance</i> 63 (2), 921–45.	381	1.776
Reinganum, Marc R. 1982. "A Direct Test of Roll's Conjecture on the Firm Size Effect." <i>The Journal of Finance</i> 37 (1), 27–35.	270	1.776
Serra, Ana Paula. 2003. "The Cross-Sectional Determinants of Returns: Evidence from Emerging Markets' Stocks." <i>Journal of Emerging Market Finance</i> 2 (2), 123–162.	31	0.021
Strong, Norman and Xinzhong G. Xu. 1997. "Explaining the Cross-Section of UK Expected Stock Returns." <i>The British Accounting Review</i> 29 (1), 1–23.	174	n/a
Tinic, Seha M. and Richard R. West. 1986. "Risk, Return, and Equilibrium: A Revisit." <i>Journal of Political Economy</i> 94 (1), 126–47.	212	4.107
Vos, Ed and Byron Pepper. 1997. "The Size and Book to Market Effects in New Zealand." <i>New Zealand Investment Analyst</i> 18, 35–45.	21	n/a
Wahlroos, Björn. 1986. "Anomalies and Equilibrium Returns in a Small Stock Market." <i>Journal of Business Research</i> 14 (5), 423–40.	16	0.007
Waszczuk, A. 2013. "A Risk-Based Explanation of Return Patterns-Evidence from the Polish Stock Market." <i>Emerging Markets Review</i> 15, 186–210.	39	0.094
Whited, Toni M. and Guojun Wu. 2006. "Financial Constraints Risk." <i>Review of Financial Studies</i> 19 (2), 531–59.	981	2.081
Zarowin, Paul. 1990. "Size, Seasonality, and Stock Market Overreaction." <i>Journal of Financial and Quantitative Analysis</i> 25 (01), 113–125.	459	0.38

Appendix B – Estimating the Mediating Factors of Publication Bias: Additional Results

Table 7: Robustness checks

	(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 (0.115)	-0.741*** (0.206)	-0.670*** (0.167)	-1.180*** (0.0653)	-1.732*** (0.0663)	-1.749*** (0.0748)	-0.0259 (0.118)	-0.835 (0.526)	-0.526*** (0.152)
SE*Impact	-0.210*** (0.0334)		-0.348*** (0.0373)	0.00863 (0.0287)		-0.101*** (0.0309)	-0.177*** (0.0320)		-0.263*** (0.0550)
SE*Pub. Year		0.0000547 (0.0000456)	0.000252*** (0.0000435)		0.000232*** (0.0000141)	0.000276*** (0.0000452)		0.0000711 (0.000102)	0.000158*** (0.0000292)
Constant	-0.0562*** (0.0202)	-0.0347 (0.0247)	-0.0396** (0.0176)	-0.000280 (0.000275)	-0.000127 (0.000246)	-0.0000701 (0.000244)	-0.0490*** (0.00824)	-0.0304 (0.0195)	-0.0377*** (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.

Appendix C – Results After Including Working Papers

Table 8: Estimating the magnitude of the selective reporting bias when working papers are included

	(1)	(2)	(3)	(4)	(5)	(6)
	OLS	FE	BE	Precision	Study	IV
SE	-0.915*** (0.0579)	-1.035*** (0.143)	-0.611*** (0.105)	-1.279*** (0.0548)	-0.657*** (0.159)	-0.533 (0.329)
Constant	-0.0312*** (0.00724)	-0.0211* (0.0121)	-0.0417* (0.0231)	-0.000592* (0.000317)	-0.0373*** (0.0125)	-0.0648** (0.0277)
Observations	1913	1913	1913	1913	1913	1909

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.

Table 9: Estimating the mediating factors of publication bias when working papers are included

	(1)	(2)	(3)	(4)	(5)	(6)
	OLS	OLS	OLS	FE	FE	FE
SE	-0.774*** (0.160)	-1.446*** (0.0840)	-1.368*** (0.109)	-1.193*** (0.187)	-1.464*** (0.204)	-1.546*** (0.224)
SE*Impact	-0.0641 (0.0545)		-0.185** (0.0726)	0.260 (0.168)		0.242 (0.156)
SE*Pub. Year		0.000176*** (0.0000122)	0.000239*** (0.0000495)		0.000151*** (0.0000455)	0.000128*** (0.0000481)
Constant	-0.0336*** (0.00682)	-0.0231*** (0.00707)	-0.0229*** (0.00575)	-0.0281* (0.0151)	-0.0165 (0.0117)	-0.0238* (0.0139)
Observations	1769	1913	1769	1769	1913	1769

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. Data on RePEc impact factor were available only for 4 out of 10 working papers added, which is why the number of observations is much smaller in specifications (1), (3), and (4). Standard errors are reported in parentheses. *, ** and *** denote significance at the 10%, 5% and 1% levels.