Does the value premium decline with investor interest in value?

Gunter Löffler*
Ulm University
This version: October 2019

Abstract

I approximate the interest that value investing attracts through the frequency with which terms such as "book to market ratio" appear in the corpus of books scanned by Google. Following the years in which investor interest in value is relatively high, the realized value premium is found to be below average. On the other hand, there is no evidence that secular trends in interest have an impact on the value premium. The results therefore do not support the hypothesis that the value effect disappears once investors have become aware of it.

Key words: value premium; investor interest; n-grams

*Institute of Finance, Ulm University, Ulm, Germany. E-mail: gunter.loeffler@uni-ulm.de

1 Introduction

The value premium is one of the strongest empirical patterns in stock returns.[1] On average, stocks with high book-to-market ratios (“value stocks”) show higher returns than stocks whose book-to-market ratios are low (“growth stocks”). When measured by the high-minus-low factor (HML) of Fama and French (1993), the average US value premium from 1926 to 2018 was 4.7% per year. In recent years, however, the realized value premium was negative, as shown in Figure 1. Over the 10 years from December 2008 to December 2018, the HML factor return was negative for seven years, with the average being −2.3%. The average return remains negative if the recession year 2009 is excluded. Could this observed decline be because of an increased amount of attention that the premium has been receiving?

Finding an answer is difficult because effectively there is just one observation: the recent 10-year dip in the value premium. Therefore, in the current paper, I study how the value premium varies with investor interest over a much longer period. Though academic researchers did not discuss the value effect until the seminal publications of the 1980s and 1990s (Rosenberg, Reid and Lanstein, 1984; Fama and French, 1992), the investment community has long regarded valuation ratios as indicators for stock selection. A well-known example is Benjamin Graham, who suggests we should examine the “ratio of price-to-book value” (quoted from Graham and Zweig, 2003, p. 349). As a proxy for investor interest in the value effect, I take the frequency with which terms such as “book-to-market ratio” or “price-to-book value” appear in the corpus of books scanned by Google. The counts are scaled by the counts for “investing” to control for general trends in related publication activity.

[1] While Fama and French (2015) have shown that the value factor is redundant once size, profitability and investment are controlled for, it still produces the largest return spread in the usual sorts on size and one additional factor (see. Fama and French, 2015, Table 1).

The data can be easily retrieved because Google has made the counts of n-grams publicly available. The linguistic term n-gram denotes a sequences of n items. For example, “book to market” is an n-gram of size three.

Briefly summarized, results are as follows: Unconditional usage frequencies of value-related terms do not predict future value premia. The data therefore do not support the hypothesis that the value effect disappears once investors have become aware of it. However, short-term variation in investor interest is significantly associated with future returns. I measure abnormal investor interest by comparing the current level of the proxy with its lagged moving average. When the abnormal investor interest metric is below the median, the average annual HML return is 9.9% compared with 1.0% when interest is high. The magnitude and significance of this effect remains if other predictors of the value premium suggested in the literature are controlled for.

Another candidate for an investor interest proxy is flows to value funds. It has two drawbacks. Available data do not extend as far back in time as the Google data. Second, the suitability of fund flows as a proxy for interest in value is doubtful because Lettau, Ludvigson and Manoel (2018) show that funds classified as value do not systematically tilt their portfolios towards stocks with a high book-to-market ratio. Nevertheless, I include fund flows as an additional control variable for the subperiod in which the flow data are available and find that this does not lower the predictive ability of the publication-based proxy.

The related literature includes McLean and Pontiff (2012), who show that the predictive ability of firm characteristics for stock returns declines after research on the characteristics has been published. Given that the authors study average effects, their results do not provide direct evidence that a particular effect, such as the value effect, fades over time. There is also a large literature on the role of investor attention. Variation in investor attention is approximated in different ways: through day-of-the-week effects (DellaVigna and Pollet, 2009); through

Google search trends (Da, Engelberg and Gao, 2011; Chen, 2017; Cheng, YiHou Huang and Hu, 2019); through news searching and reading on Bloomberg (Ben-Rephael, Da and Israelson, 2017); through events such as extreme returns that are assumed to grab attention (Barber and Odean, 2007); through events that are supposed to draw attention away (Hirshleifer, Lim and Teoh, 2009; Kempf, Manconi and Spalt, 2017); through advertising expense (Lou, 2014); through option trading volume (Wang, 2017). Some studies measure the attention of specific investor groups such as institutional investors (Ben-Rephael, Da and Israelson, 2017; Kempf, Manconi and Spalt, 2017) or local versus nonlocal investors (Cziraki, Mondria and Wu, 2019). Estimated levels of attention are used to explain trading behavior (Barber and Odean, 2007), market reaction to news (e..g. DellaVigna and Pollet, 2009), returns on individual stocks (e.g. Da, Engelberg and Gao, 2011) or on indices and commodities (Vozlyublennaia, 2014; Chen, 2017).

The present paper differs in two major ways from this literature. To the best of my knowledge, there is no study that examines investor interest in an investment approach such as value investing. In addition, information contained in Google books has not been used in this context. In the present paper, the Google search data, which many papers use and which are available from 2004, are only used for an extension of the sample period.

The literature on Google n-gram data starts with Michel et al. (2011), who illustrate how the data could be used in different fields. Pechenick, Danforth and Dodds (2015) point out that n-gram counts do not take into account the differences in how widely books are read and that the Google corpus is increasingly populated by professional texts. Although the former is relevant when making descriptive statements about the popularity of n-grams, it seems less of a problem when an n-gram is found to have a predictive ability. Barely read books will add noise to the n-gram based proxy, thus making it more difficult to find significant relationships. Regarding the second point, the inclusion of professional texts seems to be an advantage for the present paper, in which the goal is to gage the interest that investors pay to value strategies. Any trends in the proportion of professional texts should be captured by the scaling procedure used in the current paper because the n-gram used for scaling ("investing") should be more likely to appear in professional texts. If the share of these texts goes up, so should the usage frequencies of the word "investing".

A large range of papers has examined ways of predicting the value premium and the performance of other investment styles. The early literature is summarized by Asness, Friedman, Krail and Liew (2000). Chen and De Bondt (2004) suggest that past performance of a style predicts its future performance. Recent papers are partly pessimistic about the possibility to predict factor returns (Bender, Sun, Thomas and Zdorovtsov, 2018), partly optimistic (Hodges, Hogan, Peterson and Ang, 2017).

2 Data and methodology

Choice of n-grams for the investor interest proxy

The Google n-gram viewer and the downloadable underlying data provide information on the frequency with which n-grams up to a size of five occur in the corpus of books scanned by Google. The data are available on a year-by-year basis and range from the year 1505 to the year 2008. The data provided by Google also include the number of publications in which an n-gram was used at least once. I use n-gram counts rather than publication counts because otherwise, a book devoted to value strategies would receive the same weight as a publication that has a different focus and only uses one or a few value-related phrases when dealing with related literature. Because of the large number of investment texts published by UK-based publishers such as Wiley, I take the n-gram frequencies from the Google corpus of English books (version 20120701) rather than from the narrower American English corpus.

To capture investor interest in the value effect, I count the frequencies of n-grams that combine a description of a ratio, such as “book to market,” with a noun such as “value” or “ratio.” This is meant to lower the number of instances in which the counts include phrases from fields other than finance and investing, as in “we bring your book to market.” From the literature, it appears that authors writing about value use the following ratios: “book to market”; “market to book”; “price to book”; “book to price”. To identify the nouns typically used in conjunction with these ratios, I inspect 4-grams and 5-grams listed in the Google data that start with these ratios. Based on this inspection, the relevant ones appear to be “value”, “ratio” and “multiple”, as well as “equity” for the ratios “book to market” and “market to book”.

A drawback of the Google n-gram data is that they are available only up to a size of five. Therefore, I cannot determine the counts for hyphenated expressions such as “price-to-book ratio” because the hyphens are also items that increase the size of the n-gram, turning “price-to-book ratio” into a 6-gram. As hyphenated expressions are quite common especially in more recent years, I use a workaround in order not to lose relevant information. Searching in books.google.com suggests that leaving out the first noun and the first hyphen of an expression such as “price-to-book ratio” does not lead to unwanted search results. To determine the frequency of “price-to-book ratio” and “market-to-book ratio”, I therefore count the n-gram frequencies of “to-book ratio”.[2] For the other expressions, I proceed in a similar way.

My n-gram counts are case insensitive, meaning that they also pick up occurrences such as “Book to market ratio”, and I count the frequencies of possible plural forms such as “book to market ratios”. The list of n-grams used in the study is shown in Table 1, together with descriptive information about their count values from 1926 to 2008. The most frequent n-gram is “to-book ratio” with 8718 counts. Since the hyphenation of valuation ratios became only

[2] The n-gram data do not include counts for n-grams that start with a hyphen, meaning that counts for “-to-book ratio” are not available.

common from the 1980s on, usages of “to-book-ratio” are only found in 36 out the 83 years examined here. The absolute usage frequency of “price to book value” is considerably lower, but because it was used early on in the literature, its usage is recorded in 67 years. The observation that there are some years in which there are no references to valuation ratios should not come as a surprise, given that value investing has its roots in the 1920s,[3] with the seminal publication being Graham and Dodd (1934).

Having mentioned “value investing”, one could also think of taking counts for the 2-grams “value investing” or “value strategy” as a proxy for investor interest. However, usage of these expressions only picked up in the 1980s. From 1926 to 1979, “value investing” appears a mere 19 times in the Google n-gram corpus. In the same period, “value strategy” was used a bit more often (78 times), but a search in Google books shows that most early uses of “value strategy” were not related to stock market investments.

In the Google n-gram viewer, the frequency of n-grams is expressed as the absolute count divided by the overall counts of n-grams of the same size. This controls for a general increase in the number and length of the published works. However, it seems appropriate to also control for the publication activity in the relevant field. If the relative number of publications about investing increases because of a general increase in interest in investing, references to value investing could also increase relative to the entire corpus but without a corresponding increase in the interest that value investing attracts compared to other investment approaches. In the base case, I therefore scale the frequency of the value-related counts with the frequency of references to “investing”. “Investing” is obviously closely related to writing about value investing but it is also used in conjunction with other investment approaches and with general issues arising in the context of stock market investments. Of course, “investing” is also used in

[3] For a short summary of the history of value investing, see https://www8.gsb.columbia.edu/valueinvesting/about/history

a corporate context and with respect to non-equity investments, which can lower its suitability for scaling. Given that the context will still be mostly a finance one, however, this seems less critical than using more narrow candidates such as “security analysis” or “investment analysis”. At least in the early days of security analysis – this is the title of the pioneering book by Graham and Dodd (1934) – it was largely synonymous with fundamental, value-based analysis, meaning that using these expressions for scaling would tend to mask existing variation in investor interest. In any case, other choices for scaling will be considered as part of the sensitivity analysis.

Figure 2 depicts usage frequencies over time. Scaled total frequencies for the set of n-grams listed in Table 1 are shown in Panel a). There is a marked increase in usage counts from the 1980s on, which is what one would expect given that academic research on the value effect started during that time. However, there is another increase in usage counts staring in the late 1940s. From then through the 1960s, scaled frequencies are often higher than in the 1970s. In addition to these long-term shifts, there are considerable short-term fluctuations. The average difference between two local peaks is 2.86 years.

The raw, unscaled counts of value-related n-grams are also shown in Panel a). They are dominated by an exponential upward trend. As visible in Panel b), the same applies to the raw counts of all n-grams in the Google corpus and to counts for “investing”, which is the n-gram used for scaling. The latter increase more strongly than the ones for all n-grams, consistent with the observation by Pechenick, Danforth and Dodds (2015) that the share of professional texts in the Google corpus grows over time.

Specifying the empirical regression model

In the empirical part of the current paper, I examine how usage frequency is linked to realizations of the value premium. I follow a large amount of literature and capture the differences between returns of value and growth stocks through the High-Minus-Low factor (HML) of Fama and French (1993).

Given that the n-gram counts are available on an annual basis, I study annual *HML* returns by running least squares regressions of the type

$$HML_t = \alpha + \beta\, x_{t-1} + \gamma' C_{t-1} + u_t \tag{1}$$

where x represents an n-gram based proxy for the interest that value attracts, and C is a matrix of possible control variables. Entering the usage frequencies as shown in Figure 2 will mean that the x variable may have a unit root. (An augmented Dickey-Fuller tests with three lags leads to a test statistic of 0.496, not sufficient to reject the hypothesis of a unit root.) This can lead to econometric difficulties. In particular, the t-statistics could be inflated. Since the coefficients β turn out to be insignificant at usual levels in these regressions, I nevertheless show the results because the problem of finding spurious relationships then does not cast doubt on the interpretation.

To capture short- and medium-term fluctuations in the magnitude of interest in value, I compare the usage frequency to a rolling moving average to determine an abnormal usage frequency AF_t. Let F_t denote the scaled usage frequency of the set of n-grams in year t; then, the abnormal usage frequency is defined as follows:

$$AF_{t,M} = F_t - \frac{1}{M}\sum_{i=1}^{M} F_{t-i} \tag{2}$$

As noted above, the average distance from one local peak of *F* to the next is about three years. In the base case, I compute the moving average over three years, that is, one cycle, to make the moving average a reliable indicator of recent usage frequency. This choice will be varied as a part of the sensitivity analysis. Different to the unconditional usage frequency, the abnormal frequency *AF* is stationary. The Dickey-Fuller test augmented with three lags yields a statistic of −5.934 for M=3.

As for control variables *C*, I use variables highlighted in the literature. Bender, Sun, Thomas and Zdorovtsov (2018) examine 40 different indicators of macroeconomic conditions, market conditions, and sentiment. I consider the three indicators that in their study exhibit a correlation with the one-year value premium that is above 30% in absolute terms. These are the dividend yield of the aggregate stock market, the one-year inflation rate, and the personal savings rate. In addition, I examine momentum in the value effect (Chen and De Bondt, 2004), and the value spread (Asness, Friedman, Krail and Liew, 2000). The latter is the difference between the average book-to-market ratios of value and growth stocks. I represent it through the differences in the book-to-market ratios of the portfolios that Fama and French (1993) use to construct the HML factor. Momentum is captured through the lagged one-year HML return. Data on HML and the book-to-market ratios of the underlying portfolios are obtained from the web pages of Ken French[4]; the aggregate dividend is constructed from the stock market predictor data file available on the web pages of Amit Goyal.[5] For the inflation rate, I use the all urban CPI series CUUR0000SA0R from FRED available at fred.stlouisfed.org. The savings rate from 1929 to 2018 is A072RC1A156NBEA, again from FRED; savings rates for 1926 to 1928 are calculated using information in United States Bureau of the Census (1975, Tables F6-9 and F543).

As to the sample period used for the estimation, the earliest possible starting date is 1927, when the HML series begins. Since I use the one-year lag of HML as a control variable, I let the regressions (1) start in t=1928.

[4] http://mba.tuck.dartmouth.edu/pages/faculty/ken.french/data_library.html
[5] http://www.hec.unil.ch/agoyal/

3 The empirical relationship between value premia and interest in value

Main findings

I start by examining whether variation in the unconditional frequency of value-related n-grams predicts future HML returns. In regression (I) in Table 2, the scaled frequency of the value n-grams is the only explanatory variable. The coefficient is negative but not significant. The White (1980) heteroscedasticity-robust t-statistic is −0.87.[6] Once the five control variables are included, the coefficient turns positive and statistical significance declines further. Recall that the nonstationarity of the unconditional usage frequency would add further doubts to the interpretation that usage frequency predicts HML. The control variables, by contrast, help to explain variation in the value spread. Two of the controls, the value spread and the savings rate, are significant at a level better than 1%, and including the controls makes the adjusted R^2 increase from a negative value to 11%.

When the frequency of value-related expressions is entered as the abnormal frequency, its coefficients are negative, larger in absolute size and significant at a level better than 1%. This also holds when the five control variables are included. Neither the coefficient nor the statistical significance change noticeably. Following years in which the abnormal usage frequency is high, returns of value stocks are low compared to ones of growth stocks.

Regarding the economic magnitude of the effect, the standard deviation of the abnormal usage frequency is 0.23%. Therefore, a one-standard deviation change is predicted to lead to a difference in HML returns that is around four percentage points (18.01 or 17.05 times 0.23%). As an additional illustration I examine average HML returns for years when abnormal n-gram usage frequency is above its median. In these years, the average HML return is 1.0%, compared

[6] The Durbin-Watson coefficients of the regressions (I) to (IV) are in the range of 1.79 to 2.04, indicating that there are no problems with autocorrelation.

with 9.9% when the abnormal usage frequency is below its median. For the 10% of the years that have the highest investor interest, the average value premium is slightly negative at -0.3%, but not significantly different from zero (t-statistic=-0.06).

Although the results are not sufficient to establish a causal relationship, the findings are consistent with the following interpretation: increased interest in value drives up the prices of value stocks, thus driving down returns. These patterns are of a short-term nature. As the insignificance of the unconditional usage frequency shows, secular variation in investor interest to value does not predict the value premium.

Stability of the results

Given that abnormal n-gram usage frequency predicts the value premia while the unconditional frequency does not, it is important to check whether the results are robust to how abnormal usage frequency is computed. Recall that the comparison to a three-year moving average was based on the average cycle length in usage frequency. In this variation, I consider other choices for M, the number of years over which the moving averages are computed. Specifically, I set M equal to 1 and 2 and then also 6, 9, or 12, corresponding to 2, 3, or 4 cycles of a typical length. Table 3 reports the coefficients if regression (III) of Table 2 is rerun with the different choices of M. As seen in the table, the estimated effects are rather stable. Until M=9, they remain significant at the 5% level. The base case choice M=3 leads to relatively strong results but not to the strongest ones. Setting M=1 leads to a higher p-value and a higher economic significance. The estimated coefficient is smaller than with M=3, but choosing M=1 also increases the standard deviation of the proxy, from 0.23% to 0.29%, which more than offsets the lower coefficient.

Next, I examine whether the coefficients of abnormal n-gram usage frequency are stable over time. I perform rolling regressions using a 25-year window. 25 years are chosen to have a sufficiently large number of observations; the choice implies that for three of the examined sub-periods there is no overlap in the dependent variable.

Figure 3 shows the estimated coefficients together with a 95% confidence interval. For each of the estimation windows, the estimated coefficients are negative. Some of them are statistically significant at a 5% level even though the smaller sample size is likely to reduce power. Significant coefficients are also observed for recent time periods, and there is no indication that the strength of the relationship between the investor interest proxy and value premia has decreased over time.

Another way of checking stability as well as economic significance is to examine the out-of-sample success of a trading strategy that uses return predictions from a regression model. I examine mean-variance optimal portfolios when the set of assets is given by the market portfolio, the risk-free asset and HML. Data for the CRSP market portfolio and the risk-free rate are taken from the web page of Ken French. Optimal weights w without short-sale constraints are given by:

$$w = \Sigma^{-1} m / \gamma \tag{3}$$

where γ is the risk aversion coefficient, Σ is the covariance matrix and m is the vector of expected excess returns of the risky assets. For each strategy, I set $\gamma = 3$ and recursively estimate Σ and the expected return on the market with sample estimators that use data from 1927 to the time an optimization is performed. Expected HML returns are also predicted recursively, using one of the following three ways:

(i) Expected HML returns are estimated with trailing historical means.

(ii) HML returns are predicted with recursive regressions of HML on the abnormal investor interest proxy.

(iii) HML returns are predicted with recursive regressions of HML on the five control variables used in Table 2.

I determine the first optimal portfolio in 1952, the year in which the predictive regressions have 25 observations. From then until 2008, optimal portfolio weights are determined each year with new recursive estimates. Figure 4 shows the portfolio values that obtain when the three strategies are implemented in this way, with a starting value of 100 in 1952. The legend also states the Sharpe ratio of the strategy returns.

Strategy (ii), the one built on the abnormal investor interest, leads to the highest portfolio value and the highest Sharpe ratio (0.67). Using predictions from a model with the control variables leads to the lowest performance. The results thus confirm the stability of the relationship between the investor interest proxy and the value premium. The observation that other predictors do not perform reliably is consistent with the findings of Bender, Sun, Thomas and Zdorovtsov (2018).

The next robustness check varies the scaling procedure. Instead of dividing the raw counts of value-related n-grams by the counts for “investing” I consider the following alternatives: (i) divide the counts for each n-gram by the counts for all equally sized n-grams in the Google corpus, and then sum over these scaled counts; (ii) divide the raw count total by the counts for “stock market”, “stockmarket” and “stock return” as well as their plural forms. The motivation for (i) is that this corresponds to the default method used by Google; the motivation for (ii) is similar to the one given for the base case choice of “investing”, i.e. not just capture the overall trend in publication activity but also the one in the specific field of the analysis.

Applied to regression (III) of Table 2, scaling method (i) leads to a negative coefficient with a t-statistic of −1.99; method (ii) leads to a negative coefficient with a t-statistic of −2.48. For both cases, economic significance is similar to the one that obtains when I use “investing” for scaling. A one-standard deviation change in the abnormal usage frequency leads to a predicted change in the value premium of 3.45% and 3.36% for method (i) and method (ii), respectively.

As a final robustness check, I consider additional control variables. Following the literature, I included the one-year lag of the value premium but one might surmise that there are also long-term reversal effects at work. Also, the results by Bender, Sun, Thomas and Zdorovtsov (2018) did not highlight two frequently used indicators for economic and market conditions, namely the term spread and the default spread. In this robustness check, I therefore include (i) the cumulative HML return over years t-5 to t-2;[7] (ii) the term spread, defined as the difference between the long term yield of US treasury bonds and the T-bill rate; (iii) the default spread, defined as the yield on Baa-rated bonds minus the yield on Aaa-rated bonds. Data for the term spread and the default spread are taken from the data file available on the web pages of Amit Goyal. Adding these three control variables to regression (IV) of Table 2 slightly strengthens the results. The coefficient for the investor interest proxy is now −17.74 with a t-statistic of −2.84.

Hodges, Hogan, Peterson and Ang (2017) do not present the full details of their style return prediction model. Since they highlight business cycle indicators and their direction of change, I examine the Chicago Fed National Activity Index mentioned in their paper. It is available in FRED from 1967 on. Adding the index level as well as the sign of the annual change in the

[7] Additional lags of the value premium should also alleviate potential concerns that past premia may not just predict future premia but also affect future investor interest, thus leading to problems of endogeneity. Related to that, results available on request show that investor interest does not significantly depend on past value premia.

index to regression specification (IV) of Table 2 does not reduce the magnitude and significance of the investor interest proxy.

Updating the analysis

The n-gram data provided by Google end in 2008. One could think of updating the data through manual searches in books.google.com that are confined to specific calendar years. Whether this would lead to reliable figures is doubtful. For several years in which the Google n-gram data contain nonzero counts, I failed to find a hit in books.google.com. Apparently, it is not possible to reconstruct the n-gram data by conducting a search in Google books. Also, there is no straightforward way of obtaining the usage counts from the list of books that a search in Google books produces.

Instead, I consider another dataset provided by Google: Google Trends. For the terms listed in Table 1, I download the monthly U.S. search frequency data and average them for each calendar year. Search terms are entered in quotation marks to get results for the exact phrase. Since the search frequency data are normalized by Google and thus not directly comparable across search terms, I download results that Google Trends produces when searching not just for one term but also for another term for comparison. For the latter, I always choose “price to book ratio”, which in joint searches yields higher search frequency values than any other of the terms. I then average the search frequencies across terms and scale the resulting values with the search frequency data for “investing”.

The search data are available from 2004 on, meaning that there is an insufficient number of observations to check only with these data whether the search frequencies are associated with the value premium.[8] Therefore, starting with the year 2009, I chain the n-gram usage frequency and the Google Trends data. This chaining is done such that a 1% increase in Google Trends searches leads to a 1% change in the chained series. After chaining together the series, abnormal investor interest is determined as in the base case with a three-year moving average.

Regression results with the new data are shown in Table 4.[9] Columns (I) and (II) show the results when the previous specification is run on the sample extended with the help of the Google search data. They do not differ noticeably from the ones reported previously in Table 2. Higher abnormal investor interest significantly predicts lower value premia. Since the years added to the original sample are also the ones in which the value premium was historically low, it is interesting to have a more detailed look at the data. Figure 5 plots the value premium against the prior year's abnormal investor interest, distinguishing the 1928-2009 observations using Google n-gram data from the observations added with the Google search data. As is evident from the separate regression lines shown in the graph, the relationship over the years 2010 to 2018 is flatter but still negative. In the graph, I also highlight the first two years of the added years, i.e., 2010 and 2011. The regression line would look different if one excluded them, but it would still have a (much more) negative slope.

With the extended data, it is now also interesting to run the regressions for the subset of data for which mutual fund flow data are available. I obtain information on fund flows from the Lipper database available in Reuters Eikon. I define the relative flow to value funds as

$$Fund\ flow\ (value - growth)_t = \frac{Net\ flows\ to\ value\ funds_t - Net\ flows\ to\ growth\ funds_t}{Market\ value\ of\ HML\ portfolios_{t-1}} \qquad (4)$$

---

[8] Using unconditional search frequencies, the number of annual observations would be 14. If one followed the approach used for n-grams and determined abnormal search frequencies through a comparison with the trailing three-year average, the number of observations would go down to eleven.

[9] For the sake of brevity, Table 4 does not report regressions with the unconditional investor interest as a predictor. It continues to be insignificant at the usual levels.

for which I consider funds domiciled in the US and use the “US Mutual Fund Classification” scheme. “Value funds” include funds classified as Large-Cap Value, Mid-Cap Value, Small-Cap Value and Multi-Cap Value. Growth funds are selected accordingly. The market value of the portfolios entering HML can be determined from the data available on Ken French’s web page. It seems an appropriate choice for scaling because the potential price impact of fund flows will depend on how large they are relative to the universe of assets that the funds invest in.

Results are also shown in Table 4. As the Lipper start in 1992,[10] the first year for which the value premium can be predicted with fund flow information is 1993. Including fund flows as an additional control variable therefore leads to a marked drop in the sample size. This provides a possible explanation for the observation in regressions (III) and (IV) that significance values are mostly lower when the sample period is 1993–2018 rather than 1928–2018 as in regressions (I) and (II). The coefficients of the abnormal investor interest proxy are still negative but no longer significant. The same is true for the fund flow variable. Its negative coefficient is consistent with the same interpretation as the one given for the n-gram based investor interest proxy: high flows to value funds support the prices of value stocks, thus lowering future returns.

Interestingly, when the control variables including the fund flow variable are added in regression (IV), the coefficient of the abnormal interest variable and its t-statistic increase in absolute terms relative to regression (III). Controlling for fund flows tends to increase the predictive power of the n-gram based proxy rather than decrease it. This should reduce possible

[10] Actually, Reuters Eikon also displays data before 1992. However, the values are so small that they do not seem reliable. For example, the reported net asset values of growth funds as well as all equity funds increase by a factor of five from 1991 to 1992. Since Reuters states that Lipper started tracking weekly flows in 1992 (cf. https://www.refinitiv.com/content/dam/marketing/en_us/documents/fact-sheets/lipper-us-fund-flows-fact-sheet.pdf), I do not use pre-1992 data.

concerns that the investor interest proxy suggested in this paper is a poor proxy that could easily be improved.

4 Discussion: Investor interest and market efficiency

Why the average return on value stocks exceeds that of growth stocks is the subject of heated debate. Some have argued that the value premium rationally reflects risk, while others view it as the result of irrational valuations. Recent research both strengthens and weakens the rational view. Bai et al. (2019), for example, show that the book-to-market ratio can pick up rationally priced disaster risk if this is not properly modeled in the benchmark asset pricing model. On the other hand, Asness, Moskowitz and Pedersen (2013) document that value and momentum returns are correlated, which questions the rational view because there is no obvious way of explaining momentum through risk.

If the value premium is because of risk, it should continue to exist as long as risk and risk preferences do not change. If the difference is because of mispricing, it might disappear once investors are fully aware of it.[11]

The main finding of this study is that short-term variation in investor interest in value is associated with future value premia. The direction is consistent with the interpretation that an increase in investor interest leads to an increase in prices and thus to lower returns, and vice versa. In order to relate this observation to the origin of value premia, it is important to separate the long-run level of the value effect from its short-run variation.

As mentioned above, the value premium is positive even if investor interest is above the median. Hence, there seems to be a base premium, on top of which we observe variation linked

[11] See Asness, Frazzini, Israel and Moskowitz (2015) for a detailed discussion of how the stability of value premium is related to its reasons.

to investor interest. The base effect could be rational as well as irrational, and so could the variation around the base.

It is not obvious how the results of this paper can inform the debate about the nature of the base effect, because it is the effect that remains after controlling for investor interest. Regarding the documented sizeable variation of the value effect, plausibility arguments appear to favor the interpretation that irrationality plays a role. The rational explanation would require that there is more publication activity when expected premia are low. This is not necessarily what one would expect, given that low premia are likely to appear less interesting to readers than high ones. Anecdotal evidence also indicates that irrationality comes into play. The publication of Fama and French (1992), for example, advocates for a market efficiency explanation for the value effect and has led to more publications on the value premium. However, neither Fama and French nor much of the follow-up literature deal with the question of whether the value premium would be lower (or higher) in the near future.

5 Summary

In this paper, I have studied the relationship between the value premium and the frequency with which expressions such as "price to book value" appear in books. Although the long-run changes in this proxy for investor interest in value do not predict future value premia, the short-run deviations do. The past 90 years are characterized by swings in investor interest that also appear in the value premia. Higher interest is associated with lower future returns, consistent with the interpretation that an increase in investor interest leads to an increase in prices and thus to lower returns. Note that the predictive power of short-term changes in investor interest is also relevant for the interpretation of the finding that long-term changes in interest do not predict the premium. Without the former, one could easily attribute the failure to find predictive power for long-term changes to an inadequacy of the investor interest proxy suggested in this paper.

Taken together, the results point to a possibility that has not received much attention in the literature. The usual hypothesis is that mispricing disappears once investors have learned about it (see McLean and Pontiff, 2012). The patterns documented in the current paper do not support this hypothesis for the value premium. Investor interest in value has trended up over the last 90 years, without a concomitant decrease in the premium. Therefore, the literature studying how mispricing is related to the market's awareness of it is likely to benefit from a consideration of how awareness changes over time. In this study of the value premium, at least, transitory changes in investor interest are strongly associated with temporary fluctuations in returns.

References

Asness, C., Frazzini, A., Israel, R., and Moskowitz, T. 2014. Fact, fiction, and momentum investing. The Journal of Portfolio Management 40, 75–92.

Asness, C., Friedman, J., Krail, R. and Liew, J., 2000. Style timing: Value versus growth. Journal of Portfolio Management 26, 50–60.

Asness, C., Moskowitz, T., and Pedersen, L. H., 2013. Value and momentum everywhere. Journal of Finance 68, 929–985.

Bai, H., Hou, K., Kung, H., Li E., and Zhang, L., 2019. The CAPM strikes back. An equilibrium model with disasters. Journal of Financial Economics 131, 269–298.

Barber, B. M., and Odean, T., 2007. All that glitters: The effect of attention and news on the buying behavior of individual and institutional investors. Review of Financial Studies 21, 785–818.

Ben-Rephael, A., Da, Z., and Israelsen, R. D., 2017. It depends on where you search: Institutional investor attention and underreaction to news. Review of Financial Studies 30, 3009–3047.

Bender, J., Sun, X., Thomas, R. and Zdorovtsov, V., 2018. The promises and pitfalls of factor timing. The Journal of Portfolio Management 44, 79–92.

Chen, H.L. and De Bondt, W., 2004. Style momentum within the S&P-500 index. Journal of Empirical Finance 11, 483–507.

Chen, T., 2017. Investor attention and global stock returns. Journal of Behavioral Finance 18, 358–37

Cheng, C., YiHou Huang, A., and Hu, M., 2019. Investor Attention and Stock Price Movement. Journal of Behavioral Finance 20, 294–303.

Cziraki, P., Mondria, J., and Wu, T., 2019. Asymmetric attention and stock returns. Management Science, Forthcoming.

Da, Z., Engelberg, J., and Gao, P., 2011. In search of attention. The Journal of Finance 66, 1461–1499.

DellaVigna, S., and Pollet, J. M., 2009. Investor inattention and Friday earnings announcements. Journal of Finance 64, 709–749.

Fama, E., and French, K., 1992. The cross-section of expected stock returns. Journal of Finance 47, 427–465.

Fama, E. and French, K., 1993. Common risk factors in the returns on stocks and bonds. Journal of Financial Economics 33, 3–56.

Fama, E., and French, K., 2015. A five-factor asset pricing model. Journal of Financial Economics 116, 1–22.

Graham, B., and Dodd, D., 1934. Security analysis. New York: McGraw-Hill.

Graham, B., and Zweig, J., 2003. The Intelligent Investor. The Definitive Book on Value Investing. Revised Edition. New York: Harper & Row.

Hirshleifer, D., Lim, S. S., and Teoh, S. H., 2009. Driven to distraction: Extraneous events and underreaction to earnings news. Journal of Finance 64, 2289–2325.

Hodges, P., Hogan, K., Peterson, J.R. and Ang, A., 2017. Factor timing with cross-sectional and time-series predictors. The Journal of Portfolio Management 44, 30–43.

Kempf, E., Manconi, A., and Spalt O., 2017. Distracted shareholders and corporate actions Review of Financial Studies 30, 1660–1695.

Lettau, M., Ludvigson, S. C. and Manoel, P., 2018. Characteristics of mutual fund portfolios: Where are the value funds? (No. w25381). National Bureau of Economic Research.

Lou, D., 2014. Attracting investor attention through advertising. Review of Financial Studies 27, 1797-1829.

McLean, R. D., and Pontiff, J., 2016. Does academic research destroy stock return predictability? Journal of Finance 71, 5–32.

Michel, J.B., Shen, Y.K., Aiden, A.P., Veres, A., Gray, M.K., Pickett, J.P., Hoiberg, D., Clancy, D., Norvig, P., Orwant, J. and Pinker, S., 2011. Quantitative analysis of culture using millions of digitized books. Science 331, 176–182.

Pechenick, E., Danforth, C. and Dodds, P., 2015. Characterizing the Google Books corpus: Strong limits to inferences of socio-cultural and linguistic evolution. PloS ONE 10, e0137041.

Rosenberg, B., Reid, K., Lanstein, R., 1984. Persuasive evidence of market inefficiency. Journal of Portfolio Management 11, 9–17.

United States Bureau of the Census, 1975. Historical statistics of the United States, colonial times to 1970 (No. 93). US Department of Commerce, Bureau of the Census.

Vozlyublennaia, N., 2014. Investor attention, index performance, and return predictability. Journal of Banking and Finance 41, 17–35.

Wang, X., 2017. Investor attention strategy. Journal of Behavioral Finance 18, 390–399.

White, H., 1980. A heteroskedasticity-consistent covariance matrix estimator and a direct test for heteroskedasticity. Econometrica 48, 817–838.

Table 1. Frequencies of value-related n-grams considered for the investor interest proxy

The counts presented in the table are case insensitive and are based on the Google corpus of English books (Version 20120701). Column "Total count" gives the number of occurrences in the 1926–2008 period; column "Years with count >0" gives the number of years during the same period for which the counts were larger than zero. (s, p) indicates that the counts were performed not just for the singular form stated in the first column, but also for its plural form.

	Total count	Years with count > 0
book to market ratio (s, p)	347	17
market to book ratio (s, p)	1319	32
price to book ratio (s, p)	459	29
book to market value (s, p)	257	44
market to book value (s, p)	1249	51
price to book value (s, p)	2349	67
book equity to market equity	82	14
book value to market value	291	49
market value to book value	951	56
to - book ratio (s, p)	8718	36
to - book value (s, p)	2867	35
to - book equity	170	13
to - market ratio (s, p)	2720	27
to - market value (s, p)	2275	38
to - market equity	679	19
to - price ratio (s, p)	2031	53
to - book multiple (s, p)	66	15
Searched for but zero counts: to - price value (s, p), book to price ratio (s, p), book to price value (s, p), market equity to book equity, book to market equity, to - market multiple (s, p), to - price multiple (s, p)	0	0
All n-grams listed above	26830	76

Table 2. The value premium and usage frequency of n-grams related to value

The table shows the results from regressions of the following type:

$$HML_t = \alpha + \beta\, x_{t-1} + \gamma' C_{t-1} + u_t$$

where x_{t-1} is the investor interest proxy, which is based on the usage frequency of the value-related n-grams in the Google corpus of English books. These frequencies are scaled by the frequencies of the n-gram "investing". The usage frequency is either the observed, unconditional one, or the abnormal usage frequency relative to a rolling three-year moving average. The control variables include the lagged HML return, the book-to-market spread (difference between the book-to-market ratios of the value and growth portfolios), the personal savings rate, the annual inflation rate and the dividend yield of the aggregate stock market. The sample period is t=1928 to 2009. Robust White (1980) t-statistics are in parentheses.

	Investor interest proxy			
	Unconditional n-gram usage frequency		Abnormal n-gram usage frequency	
Depvar: Value premium HML	(I)	(II)	(III)	(IV)
Investor interest proxy	-2.752	1.617	-18.015	-17.048
	(-0.87)	(0.41)	(-3.01)	(-2.70)
Momentum in HML		-0.147		-0.170
		(-1.06)		(-1.34)
Book-to-market spread in HML		0.023		0.021
		(4.18)		(4.12)
Savings rate		1.008		0.902
		(2.78)		(2.83)
Inflation rate		0.457		0.501
		(1.03)		(1.14)
Market dividend yield		-1.226		-1.500
		(-0.91)		(-1.28)
Constant	0.069	-0.053	0.062	-0.014
	(3.02)	(-0.73)	(4.03)	(-0.30)
Observations	82	82	82	82
Adj. R^2	-0.001	0.110	0.073	0.187

Table 3. Regression results when varying the moving average length used in calculating the abnormal usage frequency

The table shows the results for β in regressions of the form: $HML_t = \alpha + \beta\, x_{t-1} + u_t$, where x_{t-1} is the abnormal usage frequency of the value-related n-grams in the Google corpus of English books scaled by the frequencies of the n-gram "investing". The abnormal usage frequency is determined relative to a rolling M-year moving average, where M is varied in the table. The sample period is t=1928 to 2009. Robust White (1980) t-statistics are in parentheses.

	Moving average length *M*					
	1	2	3	6	9	12
β	-16.784	-15.831	-18.015	-17.49	-13.288	-12.318
t-stat	(-3.24)	(-2.94)	(-3.01)	(-2.48)	(-1.96)	(-1.90)

Table 4. Updating the regression analysis of value premia with Google Trends data

For the same n-grams that were used to determine an investor interest proxy through counts from Google n-gram, I obtain the search frequencies from Google Trends. The two series are chained together at the end of 2008. The regression details are the same as in Table 2, columns (III) and (IV). In regression (IV) of this table, an additional control variable that is only available from 1992 on is added. It is defined as net flows to value funds minus net flows to growth funds divided by the market value of the HML portfolios. Robust White (1980) t-statistics are in parentheses.

	Sample period for dependent variable			
	1928-2018		1993-2018	
Depvar: Value premium HML	(I)	(II)	(III)	(IV)
Abnormal investor interest	-14.151	-14.038	-12.039	-19.144
	(-2.89)	(-2.74)	(-1.42)	(-1.85)
Momentum in HML		-0.170		-0.191
		(-1.36)		(-1.03)
Book-to-market spread in HML		0.022		0.190
		(4.20)		(0.99)
Savings rate		0.941		-0.159
		(3.03)		(-0.08)
Inflation rate		0.546		4.918
		(1.27)		(1.41)
Market dividend yield		-1.066		1.557
		(-0.99)		(0.26)
Fund flows (value minus growth)				-8.372
				(-1.87)
Constant	0.052	-0.045	0.031	-0.252
	(3.53)	(-1.06)	(1.08)	(-1.11)
Observations	91	91	26	26
Adj. R^2	0.054	0.183	0.037	0.221

Figure 1. Realized annual value premia and their 10-year average

The figure shows the difference between the return on value stocks and the return on growth stocks as measured by the HML factor of Fama and French (1993). A ten-year moving average ending in the year stated on the x-axis is also shown.

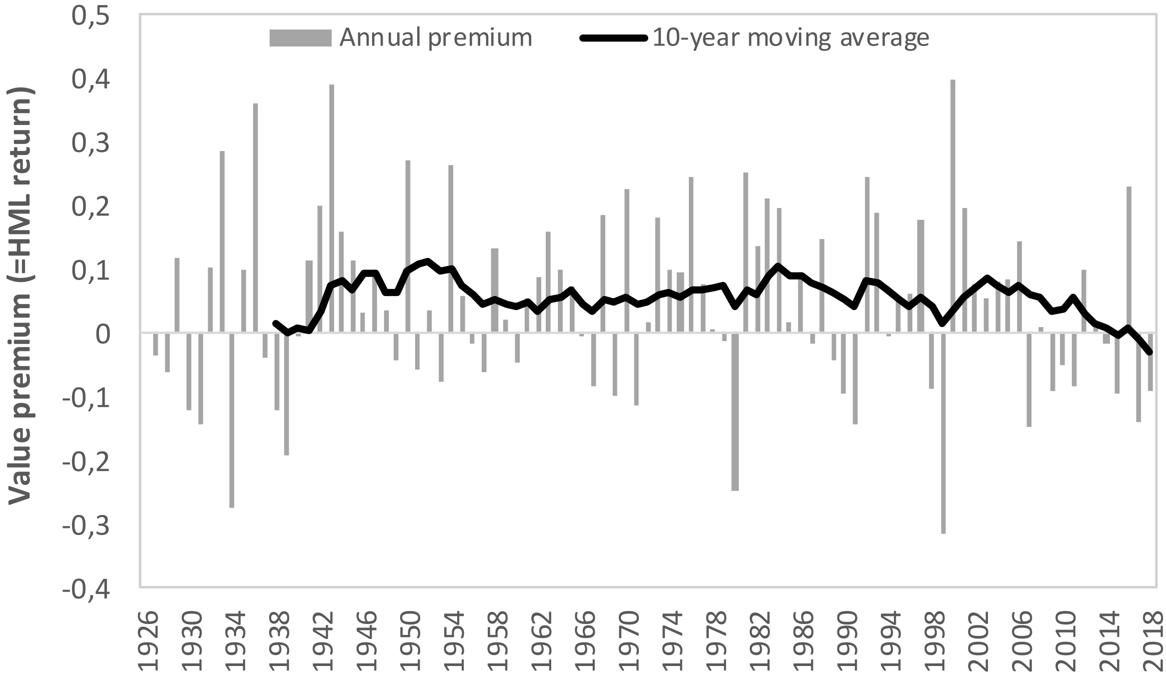
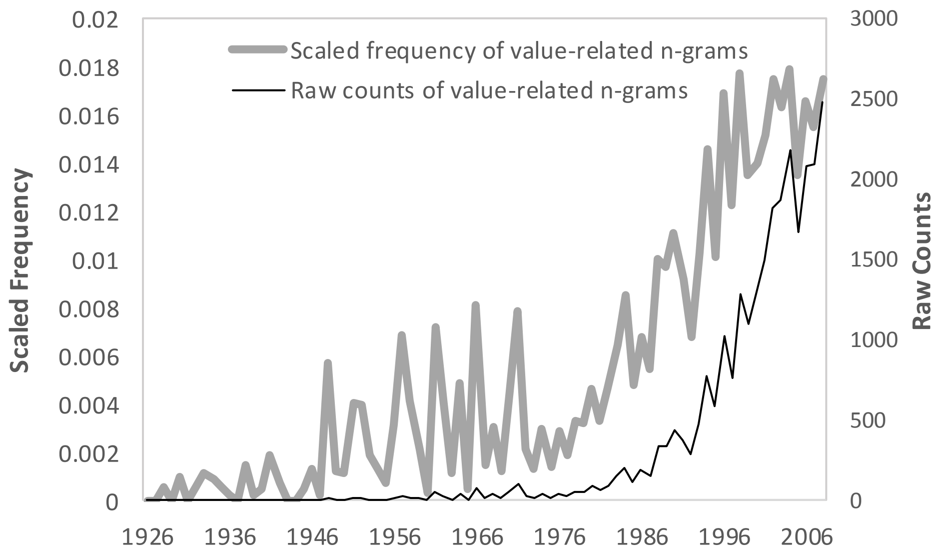


Figure 2. Usage frequency of n-grams related to value and the n-gram used for scaling

On the primary axis, the first part of the figure shows the scaled frequency with which the value-related n-grams (see Table 1 for a full list) appear in the Google corpus of books. The scaled frequencies are obtained by dividing the yearly total raw counts (secondary axis) for the value-related n-grams by the yearly counts for the 1-gram "investing". Counts for the latter and for all n-grams are show in the second part.

Panel a) Scaled Frequency and raw counts of n-grams related to value



Panel b) Raw counts for all n-grams and for "investing"

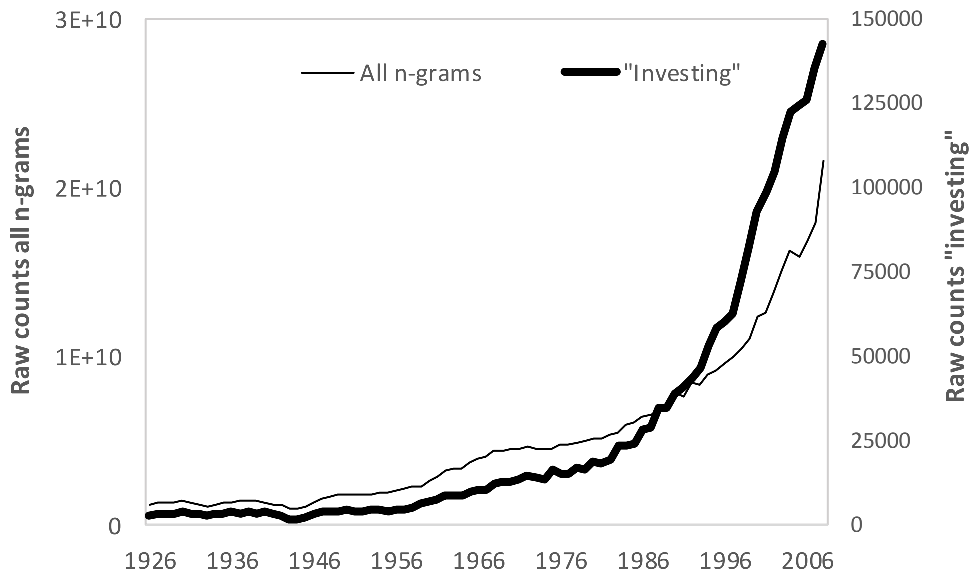


Figure 3. Stability of the regression coefficients over time

The figure shows the results for β in regressions of the following form: $HML_t = \alpha + \beta\ x_{t-1} + u_t$, where x_{t-1} is the abnormal usage frequency of value-related n-grams in the Google corpus of English books scaled by the frequencies of the n-gram "investing". The regressions are conducted for rolling 25-year windows starting in the year stated on the x-axis. Confidence intervals are based on robust White (1980) standard errors.

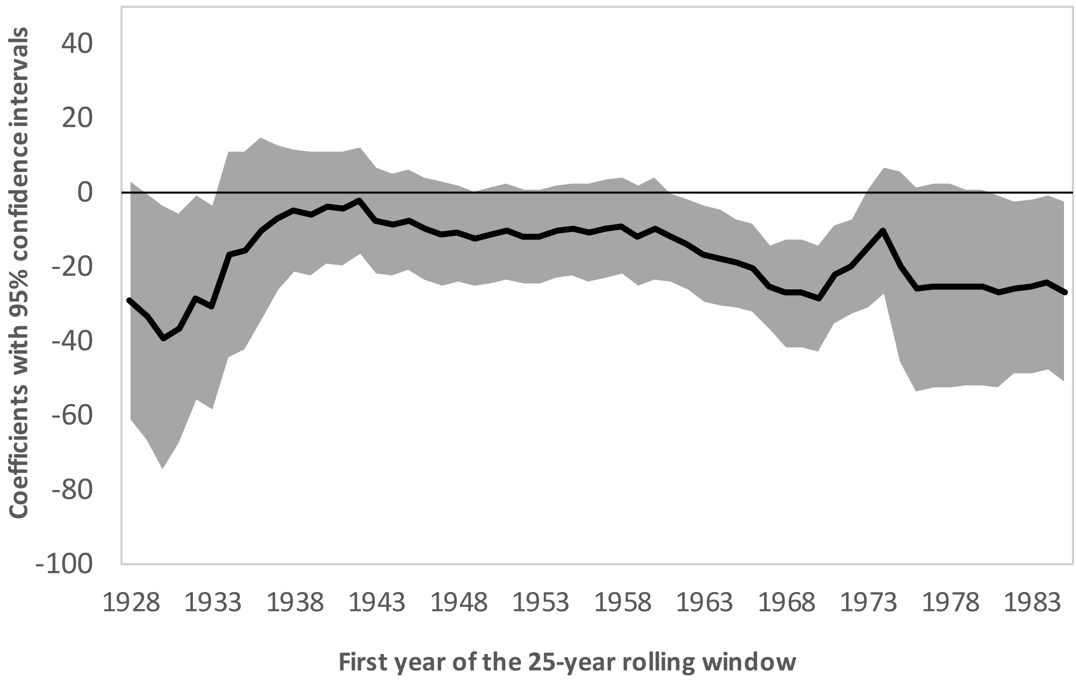


Figure 4. Performance of optimized portfolios for three different ways of estimating expected HML returns

At the end of each year from 1952 to 2008, mean-variance optimal portfolios are determined with recursive estimates of expected returns and the covariance matrix. The expanding estimation windows start in 1927. The risky assets considered are the US market portfolio and HML. Each strategy uses the recursively estimated sample covariance matrix as well as trailing historical means for the market portfolio. Expected returns for HML are estimated using (i) trailing historical means; (ii) recursive regressions of HML on the abnormal investor interest rate proxy; (iii) recursive regressions of HML on the five control variables from Table 2. Sharpe ratios of annual strategy returns are stated in the legend.

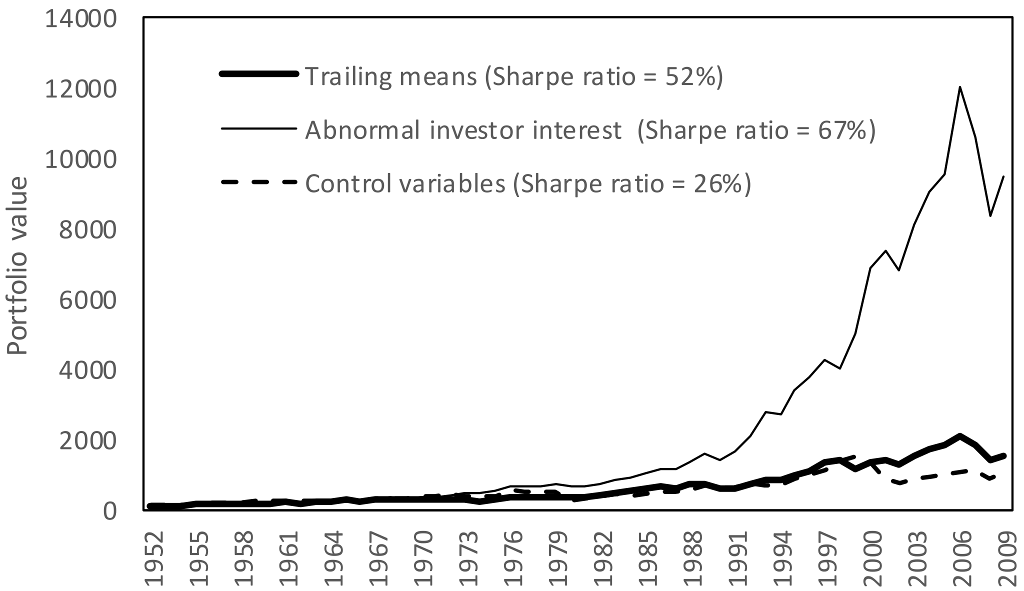


Figure 5. The value premium and the abnormal interest proxy

On the primary axis, the figure shows the abnormal investor interest proxy. Until t=2008, it is based on the scaled counts of value-related n-grams in the Google corpus of books. It is then concatenated with scaled Google search frequencies for the same value-related n-grams. Scaling is done with counts and search frequencies for “investing”. The primary axis shows the value premium as measured by the return of the HML portfolio.