J Asset Manag (2018) 19:384–393
https://doi.org/10.1057/s41260-018-0089-4

ORIGINAL ARTICLE

The impact of size and book-to-market among paired stocks

Hannes Mohrschladt[1]

Revised: 29 June 2018 / Published online: 2 August 2018

Abstract The return premiums associated with size and book-to-market also emerge between paired stocks of very similar firms. The Sharpe ratios of the traditional Fama–French-factors SMB and HML can be more than doubled if this new pair-based approach is applied. Moreover, the proposed investment strategies are particularly profitable among illiquid stocks and around earnings announcements and still yield significant premiums after 1990, while the Fama–French-factors do not. The empirical tests indicate that parts of the return premiums are due to behavioral biases indicating that investors could profit from the apparent mispricing without increasing their risk exposure.

Keywords Paired stocks · Fama–French-factors · Investment strategy · Size · Book-to-market · Behavioral finance

JEL Classification G12 · G14

Introduction

Firms with low market capitalization or high book-to-market (BM) ratios earn comparably high subsequent returns. Banz (1981) and Rosenberg et al. (1985) were among the first to report these two well-known relationships. Followed by an enormous strand of the literature, size and value effect remain heavily discussed in the financial literature even more than 30 years later. Fama and French (1993) use the observed return patterns to construct the famous Fama–French-factors SMB^{FF} and HML^{FF}. Many scholars consider the return premiums associated with SMB^{FF} and HML^{FF} to be a compensation for risk (e.g., Liew and Vassalou 2000; Petkova 2006; Elgammal and McMillan 2014). Others suspect SMB and HML to be the result of investor irrationality and mispricing (e.g., Daniel and Titman 1997; Barberis et al. 1998).

For the asset management industry, the ongoing discussion also has extensive implications, which are mainly twofold. First, the Fama–French-factors have become a widely used benchmark to evaluate and risk-adjust asset managers' portfolio performance. This application is in line with the hypothesis that the return premiums associated with SMB^{FF} and HML^{FF} are a compensation for sources of systematic risk. Second, some fund managers systematically tilt their portfolios toward small and value stocks to profit from the above-average returns potentially caused by mispricing.

In order to propose a more efficient benchmark, we aim at constructing less volatile factors *SMB* and *HML*. In order to do so, we construct pair-based Fama–French-factors that provide additional insights for both areas of application. First, abnormal returns based on size and book-to-market are even possible after controlling for the potential risk associated with the Fama–French-factors. Second, the Sharpe ratios of existing size and book-to-market strategies can be more than doubled if the new approach is applied.

Moreover, implementation is very simple: For each sample year, pairs of similar stocks are formed based on the return correlation of the previous 5 years. The small [value] stock of each pair constitutes the long position and the big [growth] stock the short position. These paired factors contain less company-characteristic noise since

✉ Hannes Mohrschladt
hannes.mohrschladt@wiwi.uni-muenster.de

1 Finance Center Münster, University of Münster, Universitätsstr. 14-16, 48143 Münster, Germany

long and short positions are very similar by construction (see, for example, Gatev et al. 2006; Do and Faff 2010; Jacobs and Weber 2015 for equivalent intuition in pairs trading strategies). Thus, the pairing procedure considerably reduces factor volatility, while the resulting pair-based factors SMB^{PB} and HML^{PB} still yield significantly positive returns. More specifically, ex ante identifiable profitable pairs yield annual return premiums beyond 10%. Since positions in the proposed size and book-to-market pair trades are held for 1 year, trading costs are comparably low. Moreover, the corresponding returns also remain significantly positive in recent years, while SMB^{FF} and HML^{FF} do not. Furthermore, time-series tests show that SMB^{PB} and HML^{PB} can explain the return premiums of SMB^{FF} and HML^{FF}, but not vice versa. This means that SMB^{PB} and HML^{PB} still yield abnormal positive returns after taking potential risks associated with SMB^{FF} and HML^{FF} into account.

Further analyses support an at least partly behavioral explanation for SMB^{PB} and HML^{PB}. First, the return premiums emerge between paired stocks of very similar firms which should have a similar risk exposure by construction. Second, the return premiums are more pronounced among illiquid stocks where limits of arbitrage are presumably high and mispricing is more likely to persist (Shleifer and Vishny 1997). Third, a disproportionate large part of the return premiums is realized around earnings announcement dates when investors seem to adjust their biased expectations. While the revision of biased expectations has become a popular explanation for value effects (La Porta et al. 1997), van Dijk (2011) notes that this kind of mispricing explanation has not yet been directly examined for the size effect.

The pair-based factors

Data

The stock return sample includes all US shares on ordinary common equity listed at NYSE, Amex, and NASDAQ from July 1965 to June 2014. For the formation of pairs, earlier return data beginning in July 1960 are used. Return and market capitalization data are obtained from the Center for Research in Security Prices (CRSP), while accounting data are sourced from COMPUSTAT (both accessed via Wharton Research Data Services). In line with the previous literature, financial firms are excluded since the interpretation of their BM ratios substantially differs in comparison with nonfinancial firms (see, for example, Fama and French 2008).

Stock return data from July of year t to June of year $t+1$ are included if the following conditions are met: At least 1 month of return data between July t and June $t+1$ and monthly returns for the previous 5 years must be available. This procedure allows for pair construction based on historical correlation coefficients and reduces potential data backfilling biases. Furthermore, the corresponding firm characteristics size and BM ratio must be available. Size is calculated as the stock's market capitalization in June t. The BM ratio is defined as the firm's book value of equity in the financial year ending in calendar year $t-1$ divided by the firm's market capitalization in December $t-1$. Consequently, the returns are measured at least 6 months after financial year end to ensure that the information has been published before the return measurement period. A firm's book value of equity is calculated as the book value of stockholders' equity plus balance sheet deferred taxes plus investment tax credit minus preferred stock book value. The book value of preferred stock is estimated using its redemption, liquidation, or par value in this order. This specific calculation procedure strictly follows Fama and French (1993) in order to benchmark the pair-based factors with their famous factors SMB and HML. The procedure results in 109,891 firm-year observations.

Construction of the pair-based factors

Previous research shows substantial variation of size and BM ratio among US firms. A substantial proportion of this variation is due to different company characteristics such as product range, business model, resource dependence, and customers (see, for example, Chou et al. 2012). But neither the risk-based nor behavioral literature assumes that these special properties are the main reason for the return premiums associated with size and BM ratio. Eliminating the return impact of these characteristics should therefore reduce noise in the measurement of the return premiums. Hence, the following procedure aims at the construction of noise-free Fama–French-factors SMB and HML. The objective to reduce variance and noise in long–short trading strategies is closely related to the search for good substitutes for individual stocks. Hedge funds use this intuition in pairs trading strategies. Based on industry and stock return characteristics, hedge funds form pairs of similar stocks. If a return spread between the two stocks emerges, one buys the loser and shorts the winner. Empirical evidence shows that the spread tends to disappear within the subsequent weeks (Gatev et al. 2006; Jacobs and Weber 2015). Consequently, the trading strategy yields positive returns on average and entails relatively low risk since the stocks in a pair are close substitutes.

We apply this intuition to construct noise-free versions of SMB and HML. Thus, the two stocks of a pair presumably share very similar characteristics such that their main difference lies in size and book-to-market. Pairs are constructed each year at the beginning of July and held for the subsequent 12 months. For each stock, its associated pair stock is determined as the stock from the same industry (identified by its two-digit SIC code) that offers the highest correlation of monthly returns for the previous 5 years.[1] If there is only one stock in an industry for a given year (0.1% of all observations), it is excluded from the analysis. The stock with the lower market capitalization [higher BM ratio] constitutes the long position, while the other stock is shorted for the following 12 months. The pair-based factors are then calculated as the equally weighted returns of each pair's long–short strategy. This procedure implies that even paired firms with very similar market capitalization or BM ratio enter the long–short strategy in the base scenario. However, the subsequent analyses also evaluate pair-based factors that exclude these pairs.

Performance of pair-based factors

The factor returns are presented in Table 1. The pair-based factor for size, SMB^{PB}, generates a significant monthly return premium of 0.29%, which exceeds the average monthly return of the corresponding Fama–French-factor SMB^{FF} by 0.03%. Since the stocks of a pair have a high return correlation, the corresponding long–short positions exhibit a comparably low volatility. Hence, the Sharpe ratio of SMB^{PB} exceeds the one of SMB^{FF} by the factor 2.56. The return premium associated with SMB^{PB} can be increased even further if only those pairs are considered that transcend a minimum difference threshold, i.e., those pairs with substantial differences in firm size between the two paired firms. Using only pairs, where one market capitalization exceeds the other one by at least the factor 100, increases the monthly return premium to 1.05%. However, these long–short strategies contain fewer pairs, are less diversified, and have higher risk, so that the Sharpe ratio remains largely unchanged.

The pair-based factor HML^{PB} constructed using the BM ratio involves an average monthly return premium of 0.28%. Note that the higher premium of HML^{FF} is due to its construction: While SMB^{FF} is based on Fama and French's entire data set, HML^{FF} is constructed as the return difference between the top 30% BM ratio firms and the bottom 30% BM ratio firms, i.e., the calculation of HML^{FF} does not include stocks with non-extreme BM ratios. Imposing a minimum factor of 1.5 between high- and low-BM firms for pair construction eliminates roughly 40% of the non-extreme pairs as well. On this comparable basis, HML^{PB} carries a premium of 0.41% which slightly exceeds the average value of HML^{FF}. Moreover, all versions of HML^{PB} clearly dominate the corresponding traditional Fama–French-factor in terms of Sharpe ratio magnitude and significance. Restricting the sample to pairs with even larger BM differences increases the monthly premium to 0.91%, which corresponds to an annual return spread of 11.5%. Given that Ready (2002) uses one-way trading costs of 13 basis points and that pair positions need to be adjusted at most once a year, transaction costs do not seem to render the proposed strategy unattractive. This also holds true, if a more conservative estimate of 80 basis points is applied as proposed by Edelen et al. (2013) who also include price impact costs.

The correlation coefficients between Fama–French-factors and their pair-based equivalents are 55.32% and 57.28% for SMB and HML, respectively. Although the pair-based factors contain effects between very similar firms only, these correlations show that both factor methodologies reflect similar information about size and BM ratio. The following subsection thus examines to which extend the profitability of the pair-based factors is captured by the Fama–French-factors.

Factor spanning tests

In order to examine whether the return premiums in Table 1 can be explained by known risk factors, we apply time-series regressions

$$SMB_t^{PB} = \alpha + \beta \cdot F_t + \varepsilon_t \text{ and } HML_t^{PB} = \alpha + \beta \cdot F_t + \varepsilon_t \tag{1}$$

where α corresponds to the intercept and ε_t to the error term. F_t is the matrix of risk factors, and β is a row vector of corresponding factor loadings. The monthly risk factors are obtained from Kenneth R. French's Web site.

The intercept and slope estimates of various factor specifications are given in Table 2. Only about one quarter of the average return premiums associated with SMB^{PB} and HML^{PB} can be explained by the corresponding Fama–French-factors SMB^{FF} and HML^{FF} in panels (1) and (6), respectively.[2] For example, the excess return of SMB^{PB} is

[1] Note that the use of correlations differs from the routinely applied distant approach in the pairs trading literature (see, for example, Gatev et al. 2006). While the distant approach measures whether two stocks followed the same overall trend in previous months, we are more interested in whether the stocks have a similar dependence on macroeconomic developments, customer movements, or input factor prices. This comovement is presumably better reflected by return correlation because stocks could be matched according to the distance approach even if their monthly returns show no common factor exposure.

[2] The other way round, empirical analyses even show that the risk premiums of SMB^{FF} and HML^{FF} can in turn be explained by their pair-based equivalents, i.e., in time-series regressions $SMB_t^{FF} = \alpha + \beta \cdot SMB_t^{PB} + \varepsilon_t$ and $HML_t^{FF} = \alpha + \beta \cdot HML_t^{PB} + \varepsilon_t$, the intercept estimates do not significantly differ from zero.

Table 1 Summary statistics for pair-based factors

	SMB^{PB}				SMB^{FF}	HML^{PB}				HML^{FF}
Min diff level	1.0	5.0	10.0	100.0		1.0	1.5	2.0	5.0	
Mean	0.0029	0.0053	0.0071	0.0105	0.0026	0.0028	0.0041	0.0052	0.0091	0.0037
SD	0.0137	0.0252	0.0312	0.0556	0.0316	0.0094	0.0139	0.0173	0.0310	0.0292
Sharpe ratio	0.7307	0.7325	0.7861	0.6518	0.2854	1.0228	1.0240	1.0488	1.0229	0.4382
t stat	5.12	5.13	5.50	4.56	2.00	7.16	7.17	7.34	7.16	3.07
Pairs per year	2240	917	566	87		2240	1324	870	199	
Corr FFfactor	0.5532	0.5781	0.5846	0.5111	1.0000	0.5728	0.5865	0.5791	0.4769	1.0000

Every year t at the beginning of July, stock pairs are constructed. For each stock, the corresponding pair stock is determined as the stock from the same industry that yields the highest correlation of monthly returns for the previous 5 years. Within each pair, the stock with the lower market capitalization in June t [the higher BM ratio in December $t-1$] constitutes the long position for the following 12 months, while the other stock constitutes the short position. The pair-based factor SMB^{PB} [HML^{PB}] is constructed as the equally weighted return of the size-sorted [BM-sorted] long–short strategy. Different versions of the pair-based factors are constructed such that a pair's return is only considered if the ratio of higher divided by lower market capitalization [higher divided by lower BM ratio] exceeds a certain minimum difference level. For comparison, summary statistics for the Fama–French-factors SMB^{FF} and HML^{FF} obtained from Kenneth R. French's Web site http://mba.tuck.dartmouth.edu/pages/faculty/ken.french/data_library.htm are presented, too. The table shows the mean, standard deviation, annualized Sharpe ratio, and t statistic for the monthly factor returns. Furthermore, the average number of pairs per year and the correlation coefficient with the corresponding Fama–French-factor is presented. The sample period covers July 1965 to June 2014

0.23% compared to its time-series average of 0.29%. Moreover, the excess returns of SMB^{PB} and HML^{PB} are also qualitatively the same if the Fama–French three-factor model (panels 2 and 7; cf. Fama and French 1993), the Carhart four-factor model (panels 3 and 8; cf. Carhart 1997), or the Fama–French five-factor model (panels 4 and 9, cf. Fama and French 2015) are used. The same holds true if factors reflecting long- and short-term reversals are considered in addition.

Robustness tests

The analyses in this subsection show that the abnormal returns associated with SMB^{PB} and HML^{PB} are also robust to a variety of different empirical specifications.

Pair construction

In the base scenario, the correlation coefficients of monthly returns of the previous 5 years are used to assign stocks into pairs. Since the return correlation structure within an industry might be subject to changes over this time period, we test whether the findings remain qualitatively the same if the pair formation is based on the daily return correlation of the previous year instead of the monthly return correlation of the previous 5 years. Panel A of Table 3 supports this hypothesis. Since a trading strategy is difficult to realize in illiquid small stocks, we also examine the paired factors after eliminating micro-caps. Excluding the bottom market capitalization decile from the analysis each year does not eliminate the return premiums. Hence, profitability persists beyond micro-caps. The reduced magnitude of SMB^{PB} is expectable since the construction of SMB^{PB} builds on differences in firm size.

Furthermore, the results are robust to different industry classification schemes (see, for example, Chou et al. 2012 for comparison of methods). The return premiums associated with SMB^{PB} [HML^{PB}] range from 0.25% [0.26%] to 0.27% [0.29%] and are all significant when pairs are formed within industries that are identified by three-digit SIC codes, four-digit SIC codes, or the 48 categories introduced by Fama and French (1997). Magnitude and significance also remain the same if pair construction is solely based on the historical return correlation without taking industry affiliation into account.

Time-series patterns of pair-based factors

We also examine the performance of SMB^{PB} and HML^{PB} in four non-overlapping subperiods with 147 return months each, i.e., July 1965–September 1977, October 1977–December 1989, January 1990–March 2002, and April 2002–June 2014. Both factors yield positive returns in all subperiods at 5% significance level. Notably, the monthly return premium associated with SMB^{PB} increases from 0.24 to 0.23% in the first two subperiods to 0.38 and 0.31% in the last two subperiods. At first glance, this increase among time is in contrast to findings by van Dijk (2011) that indicate an attenuation of the size effect in the last few decades. The return premiums associated with SMB^{FF} (and

Table 2 Factor spanning tests

	SMB^{PB}					HML^{PB}				
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
α	0.0023	0.0020	0.0021	0.0022	0.0023	0.0021	0.0018	0.0020	0.0018	0.0020
	(4.85)	(4.07)	(4.17)	(4.28)	(4.31)	(5.64)	(5.52)	(5.61)	(5.41)	(5.61)
MKT − RF		− 0.0553	− 0.0580	− 0.0561	− 0.0584		− 0.0160	− 0.0190	− 0.0141	− 0.0163
		(− 4.93)	(− 5.00)	(− 4.24)	(− 3.89)		(− 1.73)	(− 2.13)	(− 1.32)	(− 1.49)
SMB^{FF}	0.2404	0.2900	0.2901	0.2689	0.2725		0.1101	0.1102	0.1048	0.0985
	(12.38)	(14.70)	(14.58)	(11.38)	(10.98)		(9.50)	(9.17)	(8.44)	(8.45)
HML^{FF}		0.1141	0.1094	0.0986	0.0945	0.1852	0.2060	0.2007	0.1925	0.1761
		(5.16)	(4.41)	(3.75)	(3.81)	(8.70)	(13.60)	(11.74)	(9.43)	(8.63)
MOM			− 0.0144		− 0.0112			− 0.0163		− 0.0192
			(− 0.88)		(− 0.80)			(− 1.52)		(− 2.06)
RMW				− 0.0775	− 0.0773				− 0.0177	− 0.0057
				(− 2.54)	(− 2.42)				(− 0.75)	(− 0.24)
CMA				0.0302	0.0440				0.0288	0.0198
				(0.59)	(0.92)				(0.72)	(0.50)
stREV					0.0080					− 0.0091
					(0.36)					(− 0.61)
ltREV					− 0.0148					0.0348
					(− 0.52)					(1.68)
R^2	0.3061	0.4124	0.4143	0.4278	0.4299	0.3281	0.4483	0.4536	0.4524	0.4638

This table reports estimates for the time-series regressions specified in Eq. (1). SMB^{PB}, and HML^{PB} is constructed as described in Table 1. MKT − RF, SMB^{FF}, HML^{FF}, MOM, RMW, CMA, stREV, and ltREV denote the excess market return and the long–short returns associated with firm size, book-to-market, momentum (prior months 2–12), operating profitability, investments, short-term reversal (previous month), and long-term reversal (prior months 13–60). These factors are obtained from Kenneth R. French's Web site. The *t* statistics for intercept and slope coefficients are provided in parentheses and are corrected for autocorrelation up to twelve lags and heteroscedasticity following Newey and West (1987). R^2 is the coefficient of determination. The time-series regressions cover the sample period from July 1965 to June 2014 using monthly data

Table 3 Robustness summary statistics for pair-based factors

Specification	SMB^{PB}				HML^{PB}			
	Mean	SD	SR	*t* stat	mean	SD	SR	*t* stat
Base scenario	0.0029	0.0137	0.7307	5.12	0.0028	0.0094	1.0228	7.16
Panel A								
252 daily returns for pair formation	0.0024	0.0147	0.5606	3.92	0.0024	0.0091	0.9139	6.40
Excluding micro-caps	0.0018	0.0125	0.5083	3.56	0.0021	0.0087	0.8538	5.98
Industry classification: 3-digit SIC	0.0025	0.0143	0.6014	4.21	0.0026	0.0093	0.9780	6.85
Industry classification: 4-digit SIC	0.0025	0.0152	0.5656	3.96	0.0026	0.0091	0.9912	6.94
Industry classification: Fama–French	0.0027	0.0138	0.6806	4.76	0.0029	0.0091	1.0884	7.62
No industry classification made	0.0028	0.0133	0.7306	5.11	0.0026	0.0104	0.8636	6.04
Panel B								
Subperiod July 1965–September 1977	0.0024	0.0146	0.5684	1.99	0.0032	0.0102	1.0989	3.85
Subperiod October 1977–December 1989	0.0023	0.0109	0.7168	2.51	0.0025	0.0075	1.1699	4.09
Subperiod January 1990–March 2002	0.0038	0.0169	0.7778	2.72	0.0025	0.0107	0.8208	2.87
Subperiod April 2002–June 2014	0.0031	0.0117	0.9273	3.25	0.0028	0.0091	1.0790	3.78
Panel C								
January only	0.0217	0.0153	4.8974	9.90	0.0116	0.0122	3.3000	6.67
All months except January	0.0012	0.0122	0.3361	2.25	0.0020	0.0087	0.7878	5.28

This table presents summary statistics on SMB^{PB} and HML^{PB} for different methodologies and sample periods. The base scenario follows the descriptions in Table 1 and covers the entire sample period from July 1965 to June 2014. In the first specification of Panel A, the pair formation is based on the correlation of daily returns of the previous year. In the second specification, micro-caps are excluded from the analysis. The other specifications of Panel A alter or drop the industry specification used for pair formation. Panel B presents summary statistics for four subperiods of the sample. Panel C presents evidence based on all January and non-January returns. For each factor specification, mean, standard deviation, annualized Sharpe ratio, and *t* statistic are calculated based on monthly factor returns

also HML^{FF}) are indeed insignificant in the second half of our sample. Consequently, differences in size and BM ratio between very similar firms are still associated with significant return spreads, while unpaired differences are not. This lends additional support to our finding that the new pair-based factors dominate the Fama–French-factors in reflecting size and value effects.

Daniel and Titman (1997) observe that the return differences associated with size and BM ratio are substantially higher in January than during the other months. Hence, we test the pair-based factors for this kind of seasonal pattern. The average January return of SMB^{PB} equals 2.17% in comparison with an average monthly premium of 0.12% in the other 11 months. For HML^{PB}, a January effect is existent, too, but less pronounced: The corresponding average January return is 1.16% compared to 0.20% in other months. Notably however, for both SMB^{PB} and HML^{PB}, the return premiums are not only significant in January but also in the return series that exclude January.

The empirical evidence is consistent with the two most prominent January effect explanations, namely tax-loss selling and window dressing (Moller and Zilca 2008). According to the tax-loss selling hypothesis, investors sell long-term weak performers (i.e., disproportionally many small and value stocks) at the end of a year to realize losses in order to reduce tax payments. The selling pressure induces an undervaluation which is corrected in January. The data support this line of argument, since the average returns of SMB^{PB} and HML^{PB} are indeed negative in the last quarter of a calendar year. Window dressing explanations focus on the opportunistic behavior of fund managers: They might want to present well-known glamor stocks in their end-of-year fund holding reports. Consequently, institutional investors might systematically buy the big or growth stock of a pair at the end of a year and buy the small or value stock in January, which would lead to the observed pair-based factor returns.

Explaining factor performance

The presented evidence raises the question for what reason the pair-based factors are associated with significant return premiums. This distinction does not only contribute to the research debate on market efficiency, but also carries practical implications: If a considerable premium proportion is tagged as behavioral, the findings might reflect a

[3] Note that we do not use three-factor idiosyncratic volatility here because the liquidity classification would also be an indirect measure for historical factor loadings on SMB and HML then. However, the findings also hold true if idiosyncratic volatility is measured with respect to Fama and French's three-factor model.

recommendable trading strategy; otherwise, SMB^{PB} and HML^{PB} might be used to adjust portfolio returns for their risk exposure.

SMB^{PB} and HML^{PB} yield positive annual returns in 38 and 42 years of the 49-year sample period, respectively. These results indicate that the very attractive risk–return trade-off is at least partly due to mispricing in the underlying stocks. Furthermore, the similarity of the two firms within a pair also favors a behavioral explanation: Since the paired firms are very alike by construction, they should be exposed to comparable sources of systematic risk and thus should have similar expected rates of return which is in conflict to the documented persistent return predictability. Although the paired firms are very similar, the differences in size and book-to-market might still be a proxy for slightly different risk exposures. Hence, the following subsections examine more closely whether the return premiums can at least partly be attributed to behavioral mechanisms.

Liquidity

Following this behavioral line of argument, Shleifer and Vishny (1997) elaborate that an asset can deviate from its fundamental value if limits of arbitrage are high. Hence, mispricing is more likely to persist in less liquidly traded assets where arbitrageurs face high transactions costs or idiosyncratic risk which prevents them from pushing prices to their fundamental values. Consequently, the return premiums on SMB^{PB} and HML^{PB} are expected to be larger for these less liquidly traded stocks.

To test this hypothesis, we use four proxies for limits of arbitrage. Following Ali et al. (2003) and Au et al. (2009), we use idiosyncratic volatility since a high level of undiversifiable risk should render a possible arbitrage strategy unattractive for risk-averse arbitrageurs (Stambaugh et al. 2015). Idiosyncratic volatility is measured as the standard deviation of a stock's market model-adjusted returns, i.e., the residuals in an OLS regression of daily stock returns on excess market returns.[3] The second liquidity measure follows Chordia et al. (2001) and equals the dollar trading volume. Third, Amihud (2002) illiquidity is measured as absolute daily stock return divided by the corresponding daily dollar trading volume. Thus, a high value indicates that even small trading amounts are sufficient to generate large price movements. Finally, we use the high–low spread developed by Corwin and Schultz (2012) to measure bid–ask spreads as an additional source of trading costs for possible arbitrageurs.

Data for the high–low spread are directly obtained from Shane A. Corwin's Web site https://www3.nd.edu/~scorwin/. The other liquidity measures are calculated based on daily CRSP data for the previous 12 months. A stock is dropped from the sample if there are less than 200 observations available in this period. Based on this restricted sample, each stock is identified as low- or high-liquidity stock based on a median split at the beginning of July using one of the four liquidity measures introduced above. The pair-based factors are then calculated separately for the different liquidity subsamples.

Table 4 reports the returns of SMB^{PB} and HML^{PB} depending on the stocks' liquidity. For each liquidity measure, the average returns associated with SMB^{PB} and HML^{PB} are significantly larger in the low-liquidity than in the high-liquidity subsample. The return premiums associated with both SMB^{PB} and HML^{PB} are more than twice as high for the low-liquidity pairs compared to the high-liquidity pairs. While average monthly pair-based returns reach up to 0.47% for low-liquidity pairs, even for high-liquidity pairs the return premiums for SMB^{PB} and HML^{PB} remain mostly significant. Merely, the monthly premiums of SMB^{PB} are lower than 0.10% if the stocks of low Amihud illiquidity and high dollar trading volume are investigated. This finding can be attributed to the high correlation of the two variables with firm size. As a consequence, the size differential among the liquid paired stocks is substantially reduced such that the return premiums decrease. Notwithstanding, the factors' Sharpe ratios for the high-liquidity stocks are still comparable in magnitude to the Sharpe ratios of the Fama–French-factors. This is mainly due to the very low standard deviation of the pair-based factors. More specifically, all high-liquidity pair-based factors are exposed to less volatility compared to their low-liquidity counterparts. In conclusion, the results in Table 4 support a mispricing explanation for SMB^{PB} and HML^{PB} because the return premiums are higher if limits of arbitrage are more severe.[4]

[4] Although the magnitude of the return premiums is dependent on liquidity, the existence of the premiums does not seem to be driven by liquidity risk itself: Neither SMB^{PB} nor HML^{PB} load significantly on the traded liquidity factor of Pástor and Stambaugh (2003) in time-series regressions. The monthly factor data from January 1968 to June 2014 are obtained from Ľuboš Pástor's homepage http://faculty.chicagobooth.edu/lubos.pastor/research/.

[5] This line of argument is in line with a recent empirical study by Engelberg et al. (2017) who argue that increased anomaly returns around earnings announcement support behavioral explanations.

[6] Note that the Sharpe ratios from Table 5 should not be compared with those from Table 1 due to methodological differences: The Sharpe ratios in Table 5 are smaller because they are based on individual pair returns, while cross-sectional diversification effects substantially reduce the volatility (and increase the Sharpe ratios) of SMB^{PB} and HML^{PB} in Table 1. Moreover, annualized Sharpe ratios based on daily and monthly returns can differ due to autocorrelation and compound interest effects (Sharpe 1994).

Table 4 Summary statistics for pair-based factors among high- and low-liquidity stocks

	Factor	Low-liquidity stocks				High-liquidity stocks				Difference	
		Mean	SD	SR	*t* stat	Mean	SD	SR	*t* stat	Mean	*t* stat
Idiosyncratic volatility	SMB^{PB}	0.0047	0.0184	0.8920	6.24	0.0010	0.0088	0.4104	2.87	0.0037	5.62
	HML^{PB}	0.0046	0.0131	1.2082	8.46	0.0011	0.0075	0.5268	3.69	0.0034	6.93
Amihud illiquidity	SMB^{PB}	0.0037	0.0180	0.7043	4.93	0.0009	0.0121	0.2495	1.75	0.0028	3.76
	HML^{PB}	0.0036	0.0121	1.0334	7.23	0.0013	0.0090	0.4950	3.46	0.0023	4.74
Dollar trading volume	SMB^{PB}	0.0034	0.0170	0.6850	4.80	0.0007	0.0122	0.1952	1.37	0.0027	3.76
	HML^{PB}	0.0033	0.0116	0.9837	6.89	0.0012	0.0086	0.4840	3.39	0.0021	4.28
High–low spread	SMB^{PB}	0.0040	0.0184	0.7607	5.32	0.0015	0.0098	0.5322	3.73	0.0025	3.59
	HML^{PB}	0.0043	0.0129	1.149	8.04	0.0017	0.0079	0.7472	5.23	0.0026	5.15

Before pair formation, each stock is labeled as low- or high-liquidity stock based on a median split according to the liquidity measures stated in the first column. The pair formation process is carried out for each group separately yielding factors SMB^{PB} and HML^{PB} among low- and high-liquidity stocks. For each liquidity measure and each factor specification, mean, standard deviation, annualized Sharpe ratio, and *t* statistic are calculated based on monthly factor returns. The second-last column states the return difference between low- and high-liquidity pairs; the *t* statistic tests whether this difference is zero. The sample period covers July 1965 to June 2014

Table 5 Summary statistics for pair-based factors depending on earnings announcement dates

Panel A	Pair return spreads around announcements days		Pair return spreads excluding announcements days		Entire sample of daily pair return spreads	
	SMB^{PB}	HML^{PB}	SMB^{PB}	HML^{PB}	SMB^{PB}	HML^{PB}
Mean	0.000463	0.000453	0.000354	0.000190	0.000363	0.000212
SD	0.0619	0.0619	0.0521	0.0521	0.0530	0.0530
Sharpe ratio	0.1187	0.1159	0.1079	0.0580	0.1088	0.0636
t stat	7.07	8.00	13.14	9.91	13.47	11.18
Observations	2,121,018	2,121,018	23,068,441	23,068,441	25,189,459	25,189,459

Panel B	Announcement factors		Non-announcement factors		Difference	
	Mean	*t* stat	Mean	*t* stat	Mean	*t* stat
SMB^{PB}	0.000655	7.36	0.000323	12.71	0.000332	3.86
HML^{PB}	0.000649	7.64	0.000182	10.04	0.000466	5.51

The pair formation process follows the descriptions in Table 1. Panel A presents pooled summary statistics on daily long–short returns. The sample in the first two columns contains all pair returns for those dates where at least one of the two firms makes an earnings announcement around (± 1 day) that date. In the next two columns, the return of a pair is only included if none of the two firms makes an earnings announcement. The last two columns refer to the entire daily return sample. Panel A presents mean, standard deviation, annualized Sharpe ratio, *t* statistic based on clustered standard errors, and the number of daily pair returns. Panel B shows the average performance of tradable factors SMB^{PB} and HML^{PB} among those 10,807 trading days that have at least one earnings announcement in the entire cross section (± 1 day). The daily announcement factors are calculated as the average return spread of all those pairs where at least one of the two firms makes an earnings announcement around (± 1 day) that date. The non-announcement factors contain all pair return spreads without earnings announcements. The last two columns state the difference between announcement and non-announcement factors. The sample period covers July 1971 to June 2014

Earnings announcements

In standard behavioral explanations, an overvaluation of growth stocks is predominantly explained by too optimistic cash flow expectations for these stocks. This idea is in accordance with the behavioral model of biased expectation formation presented in Barberis et al. (1998) and empirically supported by La Porta (1996): Investors seem to overreact with respect to negative cash flow news, such that the stock price falls below its fundamental value accompanied by a low market capitalization and a high BM ratio. A correction of this mispricing implies higher subsequent returns for small and value stocks. These biased expectation arguments should especially apply to pairs. Comparing two very similar firms, BM differences cannot be assigned to various industry standards, but are more likely to reflect mispricing because these firms face similar growth opportunities.

Thus, the pair-based factor returns should be particularly profitable when investors realize that their cash flow expectations were too optimistic for big and growth firms. This expectation adjustment is likely to take place when new information about a firm's cash flows is released and incorporated in investors' firm value calculations. Therefore, we test whether a substantial proportion of SMB^{PB} and HML^{PB} is realized around earnings announcement days.[5]

COMPUSTAT quarterly report dates in conjunction with CRSP daily returns are used to calculate stock returns around earnings announcements. Following La Porta et al. (1997), Cohen et al. (2007) and Yan et al. (2012), 1 day before and 1 day after the announcement date are also included in the return calculations to account for possible small data errors. The daily return sample covers the period from July 1971 to June 2014 because COMPUSTAT does not provide quarterly earnings announcement data for earlier periods. A pair's daily return spread is assigned to the announcement subsample if at least one of its constituents makes an announcement on that date (± 1 day).

Panel A in Table 5 presents pooled summary statistics on pairs' daily return spreads on announcement and non-announcement days. Around earnings announcements, the average daily pair return is 4.53 basis points if long and short positions of a pair are determined by the BM ratio. On non-announcement days, the spread amounts to 1.90 basis points only. Hence, in line with La Porta et al. (1997), the value effect is considerably stronger if new information is released. Although the return volatility is higher around earnings announcements as well, the Sharpe ratio indicates a far more attractive risk–return relationship on announcement dates.[6] These effects are similar, but less pronounced for pairs formed by market capitalization: The factor premium amounts to 4.63 basis points around earnings announcements and 3.54 basis points otherwise.

While Panel A presents pooled summary statistics of individual pairs, Panel B shows that the higher return spreads around earnings announcements can also be used to construct tradable factors SMB^{PB} and HML^{PB} that carry further increased return premiums. The factors are constructed for those days (± 1 day) with at least one earnings announcement by one of the firms in the entire cross section. This procedure reduces the number of evaluated trading days between July 1971 and June 2014 from 12,333 to 10,807. Pairs' daily return spreads contribute to the announcement factors if at least one of its constituents makes an announcement on that date (± 1 day) and to the non-announcement factors otherwise. The average daily return of SMB^{PB} [HML^{PB}] is 6.55 [6.49] basis points for the announcement specification which corresponds to an annualized return of 17.94% [17.76%]. These return premiums are significantly larger compared to the corresponding daily non-announcement factors. The high abnormal performance of SMB^{PB} and HML^{PB} around earning announcements strongly supports our hypothesis that the factors are driven by mispricing which is corrected when investors adjust their biased expectations.

Hence, at least some part of SMB^{PB} and HML^{PB} can be considered as a consequence of biased expectations. Further, a risk factor explanation for the results in Table 5 seems unlikely: A particularly high expected return for a pair around report dates would imply a higher price of risk or a higher risk exposure on these specific days. But for a given date, the price of risk is the same in the entire economy and should therefore not only affect the pairs with announcement events. A higher risk exposure during report dates is doubtful, too, because these returns are mainly influenced by idiosyncratic announcement news, while it is implausible that their exposure to systematic macroeconomic risk is particularly high then. An explanation of the return premiums as compensation for idiosyncratic risk is not supported by the data either since the increase in the return premiums around earnings announcements is disproportionately larger than the increase in volatility.

Conclusion

This study combines the famous hedge fund strategy of pairs trading with the firm characteristics size and book-to-market. Trading in corresponding long–short pairs yields significantly positive returns, while volatility is comparably low. The strategy therefore incorporates both the advantage of firm similarity from traditional pairs trading and the return premiums associated with size and book-to-market. This combination implies that abnormal profits can be achieved even after controlling for the Fama–French-factors SMB^{FF} and HML^{FF}. The proposed factors SMB^{PB} and HML^{PB} have twice as high Sharpe ratios, and their premiums do not deteriorate in recent decades as opposed to the traditional Fama–French-factors. Further analyses show that the performance of the pair-based factors is particularly strong for illiquid pairs and around earnings announcements. Hence, a substantial proportion of the return premiums seems to be caused by behavioral influences and biased expectations.

These findings carry two important implications for the financial industry: First, the strategy's profitability is at least partly due to mispricing and not solely a compensation for increased risk, which shows its attractive risk–return trade-off. Second, even if investment performance is evaluated with respect to SMB^{FF} and HML^{FF}, above-average profits can be achieved. The presented results could also be considered as a starting point for the further examination and application of pair-based factors. For example, pair-based procedures might be able to boost the return premiums associated with other stock characteristics as well. These investigations would reveal whether other characteristic-based anomalies are also due to subsequent return differences between very similar firms.

Acknowledgements The author thanks Geert Bekaert, Thomas Langer, Simon Rottke, and two anonymous referees for their valuable comments.

References

Ali, A., L. Hwang, and M. Trombley. 2003. Arbitrage risk and the book-to-market anomaly. *Journal of Financial Economics* 69 (2): 355–373.

Amihud, Y. 2002. Illiquidity and stock returns: Cross-section and time-series effects. *Journal of Financial Markets* 5 (1): 31–56.

Au, A., J. Doukas, and Z. Onayev. 2009. Daily short interest, idiosyncratic risk, and stock returns. *Journal of Financial Markets* 12 (2): 290–316.

Banz, R. 1981. The relationship between return and market value of common stocks. *Journal of Financial Economics* 9 (1): 3–18.

Barberis, N., A. Shleifer, and R. Vishny. 1998. A model of investor sentiment. *Journal of Financial Economics* 49 (3): 307–343.

Carhart, M. 1997. On persistence in mutual fund performance. *Journal of Finance* 52 (1): 57–82.

Chordia, T., A. Subrahmanyam, and R. Anshuman. 2001. Trading activity and expected stock returns. *Journal of Financial Economics* 59 (1): 3–32.

Chou, P., P. Ho, and K. Ko. 2012. Do industries matter in explaining stock returns and asset-pricing anomalies? *Journal of Banking and Finance* 36 (2): 355–370.

Cohen, D., A. Dey, T. Lys, and S. Sunder. 2007. Earnings announcement premia and the limits to arbitrage. *Journal of Accounting and Economics* 43 (2): 153–180.

Corwin, S., and P. Schultz. 2012. A simple way to estimate bid-ask spreads from daily high and low prices. *Journal of Finance* 67 (2): 719–760.

Daniel, K., and S. Titman. 1997. Evidence on the characteristics of cross sectional variation in stock returns. *Journal of Finance* 52 (1): 1–33.

Do, B., and R. Faff. 2010. Does simple pairs trading still work? *Financial Analysts Journal* 66 (4): 83–95.

Edelen, R., R. Evans, and G. Kadlec. 2013. Shedding light on "invisible" costs: Trading costs and mutual fund performance. *Financial Analysts Journal* 69 (1): 33–44.

Elgammal, M., and D. McMillan. 2014. Value premium and default risk. *Journal of Asset Management* 15 (1): 48–61.

Engelberg, J., McLean, D., and Pontiff, J. 2017. Anomalies and news. *Journal of Finance*. https://doi.org/10.2139/ssrn.2631228.

Fama, E., and K. French. 1993. Common risk factors in the returns on stocks and bonds. *Journal of Financial Economics* 33 (1): 3–56.

Fama, E., and K. French. 1997. Industry costs of equity. *Journal of Financial Economics* 43 (2): 153–193.

Fama, E., and K. French. 2008. Dissecting anomalies. *Journal of Finance* 63 (4): 1653–1678.

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

Gatev, E., W. Goetzmann, and G. Rouwenhorst. 2006. Pairs trading: Performance of a relative-value arbitrage rule. *Review of Financial Studies* 19 (3): 797–827.

Jacobs, H., and M. Weber. 2015. On the determinants of pairs trading profitability. *Journal of Financial Markets* 23: 75–97.

La Porta, R. 1996. Expectations and the cross-section of stock returns. *Journal of Finance* 51 (5): 1715–1742.

La Porta, R., J. Lakonishok, A. Shleifer, and R. Vishny. 1997. Good news for value stocks: Further evidence on market efficiency. *Journal of Finance* 52 (2): 859–874.

Liew, J., and M. Vassalou. 2000. Can book-to-market, size and momentum be risk factors that predict economic growth? *Journal of Financial Economics* 57 (2): 221–245.

Moller, N., and S. Zilca. 2008. The evolution of the January effect. *Journal of Banking & Finance* 32 (3): 447–457.

Newey, W., and K. West. 1987. A simple, positive semi-definite, heteroskedasticity and autocorrelation consistent covariance matrix. *Econometrica* 55 (3): 703–708.

Pástor, L., and R. Stambaugh. 2003. Liquidity risk and expected stock returns. *Journal of Political Economy* 111 (3): 642–685.

Petkova, R. 2006. Do the Fama-French factors proxy for innovations in predictive variables? *Journal of Finance* 61 (2): 581–612.

Ready, M. 2002. Profits from technical trading rules. *Financial Management* 31 (3): 43–62.

Rosenberg, B., K. Reid, and R. Lanstein. 1985. Persuasive evidence of market inefficiency. *Journal of Portfolio Management* 11 (3): 9–16.

Sharpe, W. 1994. The sharpe ratio. *Journal of Portfolio Management* 21 (1): 49–58.

Shleifer, A., and R. Vishny. 1997. The limits of arbitrage. *Journal of Finance* 52 (1): 35–55.

Stambaugh, R., J. Yu, and Y. Yuan. 2015. Arbitrage asymmetry and the idiosyncratic volatility puzzle. *Journal of Finance* 70 (5): 1903–1948.

van Dijk, M. 2011. Is size dead? A review of the size effect in equity returns. *Journal of Banking & Finance* 35 (12): 3263–3274.

Yan, Z., Y. Zhao, W. Xu, and L. Cheng. 2012. Earnings response elasticity and post-earnings-announcement drift. *Journal of Asset Management* 13 (4): 287–305.

Hannes Mohrschladt is research assistant at the Finance Center Münster and lecturer at VWA Ostwestfalen-Lippe. After studying Finance at the University of Münster, BI Norwegian Business School, and the University of Copenhagen, he obtained an M.Sc. in Business Administration from the University of Münster. During his studies, he began his career in the financial industry in 2010 at LOYS investment company before joining the Multi Asset Group of DWS Investments in 2013.