Style Concentration in Ownership and Expected Stock Returns

Gikas A. Hardouvelis
University of Piraeus

Georgios I. Karalas
University of Piraeus

80 Karaoli& Demetriou Street, Piraeus 18534 Greece

gikas.hardouvelis@gmail.com

gkaralas@webmail.unipi.gr

September 5, 2016

Abstract:

We examine the relation of expected stock returns with fund style concentration of ownership over the period 1997-2016. Concentration is measured by the Herfindahl index of the shares of different investment styles in the ownership of stocks. Our econometric results confirm the prediction of Merton (1987) that stocks with higher concentration (lower participation) in ownership exhibit higher expected returns. The results are robust to the inclusion of known risk-factors as determinants of expected stock returns, the returns of the investment styles themeselves, plus a set of style-related control variables and other liquidity or volatility characteristics of stocks. The relation remains present over multi-year horizons of stock returns and is both economically and statistically significant.

EFM classifications: 310, 350, 370

JEL classifications: G02, G12, G23

Keywords: Asset pricing, Style Investing, Stock Ownership Concentration, Institutional Investors, Multi-period Expected Stock Returns

1. Introduction

In the last twenty years the share of stocks held by institutional investors has increased dramatically, from about 45% on average in the mid-1990s to about 80% today.[1] This large ownership makes institutions the main investor class of individual stocks today. Institutional investment behavior is, therefore, central to asset pricing. Indeed, earlier authors have provided evidence that institutional demand does affect stock prices. Gompers and Metrick (2001) found that for the period between 1980 and 1996, the increased share of institutional holdings combined with the preference of institutional investors for large companies, increased the price of large stocks. They were thus able to explain part of the disappearance of the small stock premium. Bennett et al. (2003) found evidence that increased institutional ownership can explain the increased firm-specific risk and the increased stock liquidity over the period 1983 - 1997.

A large majority of institutional investors today follow particular investment styles. An investment style is a simple rule based on some benchmarks, which enables institutional investors to reduce the number of stocks from which they construct their portfolios. There is "growth" investing, "value" investing, "index" investing, etc. Through style investing, an institutional investor concentrates on a smaller group of stocks, thus reducing his informational costs. Moreover, by being self-defined into a specific style, he or she makes it easier to be advertised and communicate his (her) services to retail investors. Finally, the style definition of a specific fund makes easier its performance measurement and evaluation, a central feature in fund management.

Although the economic meaning of the various investment styles is doubtful and the definitions of some of them appear fuzzy, the widespread use of investment styles by both retail and institutional investors is a real phenomenon that may exert a systematic impact on stock prices. It is noteworthy that at the official site of NYSE one can find the style of each stock, defined by the Style Box of Morningstar.[2]

---

[1] The upward trend of the institutional ownership begins much earlier. According to the findings of Gompers and Metrick (2001), institutional ownership on the stock market almost doubled from 1980 to 1996. Relevant evidence is also provided by Bennett et al. (2003), who report that institutional ownership was around 7% in 1950 and 28% in 1970.

[2] The relevant electronic address is the following https://www.nyse.com/listings_directory/stock. Morningstar provides analytical information about the Style Box at the following electronic address: http://www.morningstar.com/InvGlossary/morningstar_style_box.aspx

Individual investors who follow the strategy of style investing allocate their capital across different styles rather than across individual stocks. Subsequently, institutional investors follow their customers' demands and choose portfolios of stocks appropriate for the investment styles their customers wish. Individual style investors may change styles but institutional investors tend to remain stable within a class of stocks that comprise a particular investment style, as long as those stocks meet certain style criteria. For example, if mutual fund A follows the "growth" style, the fund does not change its investment strategy, but continuously holds stocks with growth characteristics. However, at the individual investor level, style investors can buy shares of mutual fund A when they want to hold "growth" style stocks and can sell its shares when they want to change style.

Style investing by institutional fund managers may end up affecting the desirability of stocks. Stocks which obey the style criteria of fund managers may become "desirable" while other stocks, which do not fit any of the criteria, may fall within the cracks and disappear from the radar screens of fund managers. Thus the daily practice of style investing can create market segmentation and a style-orientated inattention in stocks. This is because the institutional investors of each style tend follow and hold only the stocks that exhibit certain characteristics consistent with their style, and are not interested in the rest of the stocks. This kind of inattention is very similar to the one presented some thirty years ago by Robert Merton (1987). In Merton's classical asset pricing model, inattention is described as limited participation due to incomplete information about a number of stocks. Merton's model fits perfectly our context of style investing and the inattention it generates.[3]

In Merton's model, if only a small percentage of investors know about a specific stock, then when markets clear, those few investors absorb the total number of the existing supply of shares in the stock, thus moving away from their optimal portfolio. Total aggregate demand for the stock is suboptimally low, leading to a lower price than the long-run equilibrium or (in the newer terminology) "fundamental" price. Hence, in the short-run equilibrium, those few investors who chose to buy the stock end up earning a premium.[4] The

[3] Merton (1987) states that the predictions of his model are valid even if the underlying reason for limited participation is different from information incompleteness, i.e., market frictions, institutional restrictions, taxing reasons or behavioral biases, etc. See p. 488.

[4] In Merton's model, a premium also exists for the idiosyncratic volatility of the stock, as investors do not hold well-diversified portfolios anymore.

higher is the concentration of ownership on this stock, the higher is also the inattention about the stock and the lower the participation in the stock, hence the lower is its price and the higher is the premium embodied in expected returns.

Although Merton's model refers to individual investors, the predictions of the model continue to hold for style investing as well.[5] The widespread use of style investing is effectively a restriction on the behavior of institutional investors, which originates from specific customers' style demands, thus leading to varying degrees of stock inattention. We measure style inattention by the observable style concentration in the ownership of stocks. We first calculate the share of a particular style present in each stock as the sum of shares of the stock held across all institutional investors who follow the specific style, divided by the total number of shares of the stock, which are held by all institutions. We then measure the style concentration as the Herfindahl Index of the percentage shares of the investment styles in the ownership of the stocks. This index provides information about the dispersion of the ownership of the stock across the different styles. The higher it is, the higher the concentration of styles or lower their dispersion, and the higher the inattention of individual stocks.

In the empirical analysis, we explore the relation between expected stock returns and style concentration, using a time series - cross sectional quarterly panel framework from the first of quarter of 1997 to the first quarter of 2016. The quarterly frequency is dictated by the availability of stock ownership data. Our main data sources are Thomson Reuters and Bloomberg. The econometric panel analysis follows the techniques in Petersen (2009).

Our results indicate that stocks with higher style concentration of ownership earn a higher subsequent return. The unconditional annual premium for a one standard deviation difference of style concentration is 2.63% (with t-statistic 5.26), which is both statistically and economically significant. We test a variety of different specifications and in all cases the coefficient of style concentration remains significant. In the full specification case, in which we include all the control variables simultaneously, the premium for one standard deviation difference of style concentration is 2.10% (with t-statistic 2.52).

One key concern in the analysis is the possibility the effect of style concentration on expected stock returns may not originate from inattention – as Merton's model predicts - but may stem from third factors, like the strategies of the specific styles themselves. In order to

[5] See Merton (1987), p.506.

address this concern we include in our econometric analysis, first, the percentage of stock ownership by each investment style and, second, individual stock characteristics that are closely related to the determination of investment styles. In the latter case, the characteristics are the well known company size and company market-to-book ratio. They are both used as critical characteristics for the determination of the investment styles and, in addition, they are both known determinants of stock returns.

Our empirical analysis shows that after including the above set of control variables, as well as other control variables that capture well-known risk factors in the Finance literature, i.e., the market beta (CAPM), the betas of a four-factor model (Fama-French (1993), Carhart (1997)), as well as other variables such as the momentum of stocks, the idiosyncratic volatility (which appears in Merton's (1987) model), the illiquidity, the turnover, the illiquidity risk, or the leverage of each stock, the concentration measure continues to have an economically and statistically significant positive relation with subsequent stock returns.

An innovative part of our analysis is its time dimension. To examine whether the effect of style concentration is related to a dislocation from long-run equilibrium, as predicted by the model of Merton (1987), or is due to a temporary style investing effect originating from mean reversion in the sense of Barberis and Shleifer (2003), we repeat our econometric analysis using stock returns over longer horizons of 1 to 4 years ahead. At longer horizons, the magnitude of the regression coefficient relating style concentration to total cumulative multi-year returns becomes larger and is statistically significant. This evidence provides strong support that the style concentration effect is an equilibrium effect due to dislocation and is consistent with the predictions of Merton's model (1987). It is very different from the effects investigated in the literature about style investing.[6]

We also explore the robustness of the econometric relation between style-concentration and expected stock returns. First, we exclude the quarters of the financial crisis (from 2007-Q3 to 2009-Q1) and repeat the econometric analysis. The results are now even stronger, although the differences are small. Second, in order to ensure that the results are not driven by outliers, i.e. by stocks with very high style concentration, we winsorize the positively skewed concentration variable (which varies between 0.11 and its theoretical

[6] By contrast, the effects due to the shares of each individual style disappear over time i.e., specific investment-style gains or losses are transient, since they apparently depend on mean-reverting style perceptions. This evidence is consistent with the underlying theory of style investing, which is based on the original paper of Barberis and Shleifer (2003).

maximum of 1.00) at the value 0.50. The results remain similar, although now the coefficient of style concentration is higher.[7] We repeat the winsorization exercise on all independent and dependent variables and the results for regression coefficient β of the style concentration variable remain similar or become stronger.

Overall, our results provide new evidence about the effects of style investing on the price formation of stocks. The style concentration of ownership (which is equivalent to a style-related inattention and lower participation) is awarded with a return premium, which is economically significant and has a lasting feature. The results are in line with the theoretical prediction of Merton (1987) and with the empirical results of Amihud et al. (1999), who show that the effects of limited participation are present even in a stock market mainly populated by institutional investors.

The remainder of the paper is organized as follows. In Section 2 we discuss at greater length the related literature. In Section 3 we describe the formulation of our variable of style concentration in ownership and explain how it is mathematically connected with the participation variable in Merton's (1987) model. In Section 4 we describe our data and the construction of our variables. In Section 5 we provide a preliminary statistical analysis of our variables and their correlation structure, and illustrate some basic stylized facts about institutional investing and about style investing. In Section 6 we present at length the main econometric analysis of the quarterly horizon. In Section 7, we extend the analysis to multi-period horizons. In Section 8, we interpret our findings and conclude.

Appendix A contains a detailed description of the 32 investment styles used to in our econometric analysis. Appendix B contains a derivation of the relation between our stock concentration index and Merton's participation index. Appendix C provides additional econometric analysis, which investigates the robustness of the relation between ownership-concentration and expected stock returns. Appendix D shows the contrast between the more lasting effect of style concentration on multi-year stock returns vs. the temporary effects of style investing.

[7] This is a mechanical increase due to the truncation of the high values of style concentration.

2. Related literature

Our study is closely related to the theoretical paper of Merton (1987), who develops a capital market equilibrium model with incomplete information and shows that participation in the ownership of stocks (or equivalently to our framework, its inverse, the concentration of ownership of stocks to only some investors) is a determinant of stock returns, along with market beta, the size of the company, and the idiosyncratic risk of the stock. Specifically, Merton shows that limited participation leads to lack of aggregate demand for the stock and a lower price in equilibrium. This lower price is equivalent to a higher expected rate of return. Whatever the underlying reason for the under-participation, the predictions of the model remain. Our paper can be interpreted as an empirical examination of Merton's hypothesis, which uses investment styles to capture the degree of investor participation in stocks. It is the first paper, which uses institutional investor data and their investment styles as a proxy for investor inattention. The results are in line with the predictions of the Merton model.

Previous empirical evidence provides indirect support for the hypothesis that decreased participation in the ownership of a stock (either due to limited information or due to limited stock liquidity) is connected with a lower stock price and a higher expected return. Arbel et al. (1983) find that firms with less analyst coverage offer a premium as compensation for informational deficiencies. Amihud et al. (1999) find more direct evidence that a reduction of the minimum trading unit in Japanese stocks increases the number of investors who own stocks of the firm, which then leads to an increase in the stock price and a decrease in the expected return. Our paper complements this literature by providing a much more direct test of Merton's theory, yet at the level of investment styles, rather than the level of individual investors.

Our paper is also related to the branch of literature, which examines the effect of style investing on stock prices. In an influential paper, Barberis and Shleifer (2003) develop a theoretical model of style investing. The key assumption of the model is that investors move funds among styles according to their relative performance. Their model predicts excess comovement between stocks belonging to the same style, less comovement between stock belonging to different styles, a momentum effect at the style level, as well as a negative cross-correlation between the returns of "opposite" styles. The momentum effect is present in the short-

run, whereas in the long-run, the situation reverses as prices mean-revert, namely move towards their fundamental value.

The studies of Teo and Woo (2004), Froot and Teo (2008), Boyer (2010), and Wahal and Yavuz (2013) confirm the theoretical predictions of Barberis and Shleifer, using US stock data. This literature focuses on the significance of styles for the explanation of the momentum – reversals phenomena and for the stock return comovement. Compared to those papers, our paper adds the element of ownership style concentration. Our paper does corroborate the presence of mean reversion in style investing, on which the earlier literature was built on, yet it also reveals that the effect of style concentration in ownership is an extra effect on top of the effects of style investing. More importantly, the concentration effect remains present in the long-run, and is economically and statistically significant, whereas the effects of style investing are only temporary.

In a paper with a different perspective than ours, Chen et al. (2002) use the number of mutual fund owners in a stock, relative to the total number of mutual funds in their sample, to proxy how binding the short-sales constraint is. Taking into account the fact that regulations restrict mutual funds from executing short-sales, the authors use the number of mutual fund holders as a proxy of the negative opinions about a stock. Mutual funds that are pessimistic about a stock cannot sell it short, but instead they must simply stay out of it. Thus, a smaller number of mutual fund owners in a stock could mean that the stock is overpriced and would subsequently underperform stocks with a higher number of mutual fund owners. Their empirical results seem to confirm their hypothesis. At first glance, this is seemingly an opposite result to ours. However, their empirical proxy focuses only on mutual funds using their non-short-sales characteristic, while we focus on all institutional owners, using their style orientation. Their period of inquiry is 1979-1998, ending about when ours begins. But more importantly, in their sample, mutual funds only hold 8.6% of the stocks, while short sales represent less than 5% of the transactions in 98% of the stocks. In our investigation institutional investors hold the overwhelming majority of stocks. Their approach is thus only indirectly related to Merton (1987), whereas ours is a direct test of Merton. Finally, the two papers are not necessarily mutually exclusive, since style inattention could be present simultaneously with binding short-sales.

3. The Style Concentration in Stock Ownership

We calculate the style concentration (H) in the ownership of stock i (for the quarter q) as the Herfindahl index of the percentage share of each investment style s ($s = 1, \dots, S$) that is present in the stock:

$$H_{i,q} = \sum_{s=1}^{S} w_{i,q,s}^2 \qquad (1)$$

The uppercase S is the total number of the different investment styles that are present in stock i (at quarter q) and the $w_{i,q,s}$ is the percentage share of investment style s, in stock i, for quarter q:

$$w_{i,q,s} = \sum_{j=1}^{J} w_{i,q,j} \quad (2)$$

The uppercase J is the total number of funds that own stock i and follow investment style s, at quarter q. The $w_{i,q,j}$ is the percentage share of each fund j ($j = 1, \dots, J$) that is owner of stock i and follows style s, at quarter q.

Our data set does not include investors who manage portfolios with value less than $100 million. Those investors are not required to file Form 13F every quarter, the legal form which provides the basis for the construction of our main independent variable, *H*. Hence we exclude them from the analysis and concentrate only on the universe of large investors.[8] The weights in equation (2) are weights within the group of investors who file form 13F. This is the correct way to calculate *H* in the absence of information on the style of small investors.[9]

[8] Leaving the smaller investors out of the calculation of index *H*, makes the implicit assumption that those exluded investors do not cause changes in the ownewrship weights of the different styles in a stock, had they chosen a fund manager for their investing decisions. Of course, part of their style-oriented demand would be offset between them (Kumar (2009)), hence the net effect of excluded investors on the weights of the styles is even smaller.

[9] To make this point clear, consider the following example: Let us compare two companies, A and B, identical in all characteristics except for the structure of their stock ownership. In company A, two different investor styles are present, each with 30% holdings, while the remaining 40% is owned by small investors whose style is unknown. At company B, there are three different styles present, each with 30% holdings, with the remaining 10% owned by small investors whose style is unknown. It is obvious that stock A has a higher concentration of investors than stock B, since small investors do not contribute to the concentration.
Observe that our chosen strategy correctly calculates the Herfindahl index *H* to be larger for stock A. For stock A, $H = (1/2)^2 + (1/2)^2 = 1/2 = 0.5$. For stock B, $H = (1/3)^2 + (1/3)^2 + (1/3)^2 = 1/3 = 0.33$.
However, had we taken into accounts the small investors in our universe of investors when calculating the style-shares w, we would have reached a different and wrong conclusion: The Herfindahl index *H* for stock A would equal $(0.3)^2 + (0.3)^2 = 0.18$ and the *H* for stock B would equal $(0.3)^2 + (0.3)^2 + (0.3)^2 = 0.27$. This methodology would wrongly have shown that stock A has lower concentration in ownership than stock B.

Another issue of concern is the correspondence between Merton's variable of participation and our variable of style concentration. Merton defines as q the ratio of the number of investors that hold a stock (N_k) over the total number of investors in the economy (N): $q = N_k/N$. A higher q means a higher participation on the stock, or equivalently a lower concentration of its ownership. In order to facilitate the comparison between q and H we could define the reciprocal of q as the concentration of ownership: $1/q = N/N_k$. To further simplify our comparison, we eliminate the numerator N, since it is common for all the stocks, hence the (corresponding to the Merton's model) concentration variable becomes: $1/N_k$. That is, the style concentration of ownership according to Merton's model is one over the number of different investment styles that are present in the stock.

Our measure of style concentration goes beyond Merton and utilizes the relative sizes in the stakes of the different styles. This way, we relax the strict assumption of Merton that among investors present in a stock, each one holds an equal amount of the stock. In our framework, we allow investors present in the stock, to hold unequal parts of a stock. In other words, we allow cases where only few of the stock owners absorb most of its supply, while the rest hold only a small fraction of the supply.

The difference with Merton is made clear from the following example: Suppose a company A has five owners, four of whom hold 1% each and the fifth owns the rest 96% of the shares. Of course, the large shareholder absorbs almost the whole stock supply and, at the same time, moves away from his optimal portfolio.[10] In equilibrium, a return premium arises due to the increased ownership concentration. Next, suppose that another company B also has five owners, each of whom owns 20% of the shares. In the case of company B, each of the five investors moves less away from his optimal portfolio, when compared to the large shareholder of the first case. Thus a lower premium should arise relative to the first case of company A. Our Herfindahl index *H* captures the distinction between the two companies A and B and the essence of Merton's model, which has to do with investors' participation in risk sharing. *H* is higher in the first case, where most risk falls on one of the five investors, with a value of 0.92, and is lower in the second case, where risks are distributed between the five investors, with a value of 0.2. By contrast, a model in which relative shares do not

[10] This is true under the assumption that the large owner does not hold a disproportional share of the total wealth, which is a realistic assumption.

matter, would deliver the same concentration parameter of 1/5 = 0.2 in both cases and would miss a lot of the information.[11]

The Herfindahl index *H* of the investment styles is a better statistic to capture concentration than the simple number of different styles present in the ownership of a stock. This is because the total number of different styles is not very large (32 in our sample), hence it is likely the number of styles present in a stock does not vary much from stock to stock. Almost all styles are likely to be present in many of the stocks, hence in those stocks the simple number of investment styles would deliver a statistic of 100%.

Digging deeper into the meaning of the Herfindahl index *H*, it effectively measures the proximity of the style-related characteristics of a stock to their corresponding cross-sectional means.[12] The intuition is that if some of the characteristics of a stock are distinctively away (either higher or lower) from their cross-sectional mean, the stock attracts the attention of institutional investors who follow the corresponding investment style, but lacks the attention of the rest. As a result, the H index of such a stock would be higher from its cross-sectional mean. On the other hand, the opposite holds for a stock whose style-related characteristics are close to their cross-sectional means. The H index of such a stock would be lower than the relevant cross-sectional mean.

The use of Herfindahl Index is not new to the literature that examines concentration of ownership. Greenwood and Thesmar (2011) use the Herfindahl index of ownership, weighted by the volatility and the correlation of the trading needs of the investors to estimate price fragility. Barabanov and McNamara (2002) and Agarwal (2007) also use the Herfindahl Index as a measure of the concentration of ownership and study its relation with stock liquidity.

---

[11] Besides, if we assume that the stocks are equally divided to their owners (let say x value for each of the owners), then our measure equals to $1/N_k$ (the Merton's equivalent):
$H = \sum(x/MV)^2 = N_k \cdot (x^2 / MV^2) = N_k \cdot (x^2/N_k^2 \cdot x^2) = 1/N_k$. If we relax the assumption of the equal divided shares (thus x_j is the value that the investor j holds in shares of the stock), our measure equals to $1/N_k + (1/MV^2) \cdot \sum_j (s_j - \bar{s})^2$ (which is the Merton's equivalent plus a positive quantity accounting for the value concentration). The proof is provided in Appendix B.

[12] The main style-related characteristics of a stock are the size and market-to-book ratio, but investors could also see the growth rate of the EPS, the dividend yield, the price momentum and others.

4. Data Sources and Variables

Our sample begins in the first quarter of 1997 and ends in the first quarter of 2016, consisting of a total of 77 quarters. The quarterly frequency is dictated by the availability of our main independent variable, the style concentration parameter *H*, which is calculated from ownership data.[13] The sample consists of 1295 NYSE common stocks, which were actively traded in 2013. The effective number of stocks that we actually utilize in our sample varies slightly from quarter to quarter. This is because some stocks disappear or, more likely, we do not have full information for all the variables of a stock during all quarters. We also exclude quarters of stocks with negative book-to-market values and stocks for which we do not have ownership data (see Table 1 for the data availability). Note that the average number of stocks in the cross-section over the entire quarterly sample period is 927. In the econometric analysis we utilize an average of 838 stocks as some of the independent variables are missing.

4.1 Institutional Data

Data for institutional investors are from Thomson Reuters[14] and are based on the mandatory 13F filings.[15] Investors that exercise investment discretion over $100 million should report their holdings of financial assets on a quarterly basis, within 45 days of the end of the quarter for which the report is filed.[16] We have access to these data through Thomson Reuters from the first quarter of 1997 and thereafter. For each stock of our sample, we are in a position to know the number of its 13F owners and their number of shares in the stock. In addition, Thomson Financial provides information about the investment style that is followed

[13] The maximum number of quarters used in the panel analysis is 76 and not 77, as returns are measured one quarter after the quarter in which the concentration parameter *H* is observed. Also, in the panel analysis we make use of constructed variables, like pre-existing factor betas. For this reason we sometimes use stock data going back to the beginning of 1995.

[14] Through its products also called: Thomson Financial, Thomson One and Thomson Reuters

[15] U.S. Securities and Exchange Commission (SEC) provide information about 13F filings in its website: https://www.sec.gov/answers/form13f.htm

[16] The four quarters are calendar quarters, they end at March, June, September and December of each year.

by those who file, based on their portfolio characteristics.[17] The data base uses thirty two different style options for the classification of institutional investors.[18]

According to Thomson Financial: "*In classifying the dominant style of an institutional investor, Thomson Financial employs quantitative techniques based on the key financial fundamentals of the individual stocks that constitute a given portfolio. Each position is weighted by its percentage of the total assets under management for a given institution or mutual fund. For each position in a portfolio, Thomson Financial compares the fundamentals of the individual stock to that of the S&P 500 Index to determine if:*

- *The forward PE of the stock is higher or lower than the S&P 500 average*
- *The indicated dividend yield of the stock is higher or lower than the S&P 500 average*
- *The 3 to 5 year projected EPS growth rate in First Call*[19] *is higher or lower than the S&P 500 average*

By aggregating each of the individual stock selections and looking at the percentage breakdown of total assets in the categories outlined above, Thomson Financial is able to assess the interplay of growth, value, and income that drives the stock selection process of each institution and mutual fund. All three fundamentals are typically used in defining each style. To be classified in a given style, an institution must generally meet all the criteria." [20]

The techniques, which are used by Thomson Financial, are the prominent techniques of classification of funds into investment styles. Chan et al. (2002) find that both the factor loadings of a fund and its portfolio characteristics give similar results about the style classification of a fund. However, they find that the approach which is based on the portfolio characteristics, predict fund returns better.

---

[17] The investors who file are institutional investors of all sorts. In some cases, Thomson Financial classifies an institutional investor to a specific investment style not by inspection of its holdings but from its current transactions, as this may be more precise about its investment style. The exact method of this alternative way of classification is proprietary.

[18] In alphabetical order: "Aggressive Growth", "Arbitrage", "Broker-Dealer", "Capital Structure Arbitrage", "Convertible Arbitrage", "Core Growth", "Core Value" ", "CTA/Managed Futures", "Deep Value", "Distressed", "Emerging Markets", "Emerging Markets Hedge", "Equity Hedge", "Event Driven", "Fixed Income Arbitrage", "Fund of Funds Hedge", "GARP", "Global Macro Hedge", "Growth", "Hedge Fund", "Income Value", "Index", "Long / Short", "Market Neutral", "Mixed Style", "Momentum", "Multi Strategy", "Quantitative", "Sector Specific", "Specialty", "VC/Private Equity", "Yield". We report the definitions of each style at Appendix A.

[19] First Call is a Thomson First Call is a branch of Thomson Financial and it is a major provider of estimates.

[20] http://www.tfsd.com/marketing/banker_r2/HomeFAQs.asp

For the purposes of the analysis, for each stock, we sum up the number of shares of all the owners of the stock among the 13F filers, who follow the same investment style. For each of the 32 styles, we thus calculate the total number of shares that belong to the style. We then sum up the shares of the 32 styles to a grand-total of shares and calculate the frations of the grand-total belonging to each style. These fractions (which sum up to unity) are the weights used in the subsequent construction of the Herfindahl Index.

4.2 Market Data

See Table 1 for the details in the construction of the variables. Data about stock prices, share volume, market capitalization, market-to-book value and debt-to-asset ratios come from Bloomberg. We take the Fama – French factors, the momentum factor and the risk free rate from the site of Kenneth French.[21]

The main dependent variable, the stock return of quarter q, is the percentage change of the stock price from the end of the previous quarter $(q-1)$ to the end of the current quarter (q) plus the dividend yield that corresponds to quarter q. Quarterly stock returns are from Bloomberg.

We take the end-of-quarter market capitalization also from Bloomberg. Market capitalization is the product of price per share times the number of shares at the end of the quarter. We use the natural logarithm of market capitalization. The market-to-book value ratio is also provided by Bloomberg and is the ratio of price per share to the book value per share (see Table 1 for the exact timing). We use the natural logarithm of the market-to-book value ratio. The debt-to-assets ratio is also from Bloomberg. It is a measure of leverage and reflects the total debt of the company divided by its total assets. Again, we use the natural logarithm of the debt-to-assets ratio. For each of the three aforementioned variables, we use the last available value of each quarter.

The turnover is calculated as the quarterly mean of the daily ratio of the shares that are traded during each day of the quarter to the total outstanding number of shares for the corresponding day. We take the trading volume and the total number of shares from Bloomberg. With the same data we calculate Amihud's ILLIQ variable (Amihud, 2002), as the quarterly mean of the daily ratio of the absolute return (percentage price change) to the

[21] http://mba.tuck.dartmouth.edu/pages/faculty/ken.french/

dollar volume (which is the shares volume times the price of the stock).[22] ILLIQ is an illiquidity measure of price impact and is widely used in the literature. Its rationale is that if for a given level of trade there is a large price impact, the stock must be relatively illiquid. Within our sample, ILLIQ decreases on average to half its original magnitude after the first five years. For this reason, we use the cross-sectionally normalized value of ILLIQ for each quarter.[23]

We estimate the betas of a four-factor model (Fama and French (1993), Carhart (1997)), by running rolling time-series regressions (with a 24-month window) of the monthly excess stock returns to the following four factors: excess market return (Rm-Rf), SMB (small-minus-big), HML (high-minus-low) and MOM (winner-minus-losers). In addition, we estimate a measure of illiquidity risk by running rolling time series regressions (with a 24-month window) of the monthly excess returns of a stock on the innovations of market ILLIQ (measured as the cross-sectional mean of the ILLIQ values of the individual stocks).[24]

We estimate the idiosyncratic quarterly volatility of the daily stock returns for each quarter, as the standard deviation of the daily risk-adjusted returns, which are estimated as the residuals of daily time-series regressions (over the whole sample) of the excess stock returns on the 4 factors of the Carhart model. We use the natural logarithm of idiosyncratic volatility in our analysis.

We also calculate a momentum variable (Jegadeesh and Titman (1993), as the three-quarter cumulative stock return of the period which starts at the end of quarter q-4 and ends at the end of quarter q-1, hence it is observed one quarter prior to the date of the measurement of returns. We exclude the last quarter to avoid any short-term reversal effects.

---

[22] $ILLIQ_{i,q} = 1/D \cdot \sum_{d=1}^{D} |r_{i,d}| / \$volume_{i,d} \cdot 10^6$, where $r_{i,d}$ is the daily price change of stock i at day d, $\$volume_{i,d}$ is the dollar volume of stock i at day d, D is total number of trading days during the quarter q, and $10^6$is a scale factor.

[23] We estimate the normalized ILLIQ for each quarter by subtracting the cross-sectional mean of ILLIQ of that quarter and then by dividing with the cross-sectional standard deviation of that quarter: $standILLIQ_{i,q} = \frac{ILLIQ_{i,q} - \overline{ILLIQ_q}}{s.d.(ILLIQ)_q}$.

[24] We measure the innovations as the residuals of an AR(1) model. As a control we also include the excess market return series in the time series regressions. The notion of illiquidity risk is developed in the papers of Pastor and Stambaugh (2003) and Acharya and Pedersen (2005) and its rationale is that if the price of a stock is sensitive to changes in market-wide illiquidity, the stock is more risky and hence investors demand a return premium in order to hold it.

We finally calculate for each stock and each quarter the total percentage of ownership of each investment style. There are 32 such variables, which are measured across 77 quarters and across all stocks per quarter. We use them as controls for possible style effects.

Table 1: Data and Variables

The first column contains the name and notation of the variable used in the analysis, the second column its definition, the third column the data sources or the data used to estimate the variable and the fourth column the number of available observation for each variable.			
Variable	Definition	Data Source	Number of Observations
Return $ret_{i,q+1}$	The quarterly return of stock i during quarter q+1 is measured as the percentage change of the price of stock i from the end of quarter q to the end of quarter q+1, plus the dividend yield which corresponds to quarter q+1: $ret_{i,q+1} = \frac{Price_{i,q+1} - Price_{i,q}}{Price_{i,q}} + \frac{dividend_{i,q+1}}{Price_{i,q}}$	Bloomberg. (Bloomberg Datatype: DAY_TO_DAY_TOT_RETURN_GROSS_DVDS)	79,214
Style Concentration $H_{i,q}$	Style concentration for stock i at quarter q is the Herfindahl Index of the weights of each style s, present in the stock during quarter q: $H_{i,q} = \sum_{s=1}^{S} w_{i,q,s}^2$. The share of each style s is estimated as the sum of shares of stock i, held by funds which follow style s. The base for the estimation of the weights is the sum of share holdings in the 13F filings.	Thomson Reuters (or Thomson One or Thomson Eikon)	72,880
Size $\ln(mv)_{i,q}$	The natural logarithm of market capitalization of stock i at the end of quarter q.	Bloomberg. (Bloomberg Datatype: CUR_MKT_CAP)	78,751
Market-to-Book $\ln(mtb)_{i,q}$	The natural logarithm of the ratio of the market value to the book value of stock i. Market value is the market capitalization at the end of quarter q and Book value is the accounting value of the stock i at the end of the previous year.	Market-to-Book ratios are directly provided by Bloomberg. (Bloomberg Datatype: MARKET_CAPITALIZATION_TO_BV)	76,075
Price Momentum $mom_{i,q}$	The cumulative stock return measured over 3 quarters, from the end of quarter q-4 to the end of q-1: $mom_{i,q} = \frac{Price_{i,q-1} - Price_{i,q-4}}{Price_{i,q-4}}$	Prices from Bloomberg. (Bloomberg Datatype: PX_LAST)	77,672
Debt-to-Assets $\ln(dta)_{i,q}$	The natural logarithm of the ratio of total debt to total assets of stock i at the end of quarter q.	Debt-to-Assets ratios provided directly by Bloomberg. (Bloomberg Datatype: TOT_DEBT_TO_TOT_ASSET)	79,624
Share Turnover $turnover_{i,q}$	Share turnover of stock i for quarter q is the quarterly average of the daily ratios of the number of shares traded each day of the quarter to the total outstanding number of shares each day of the quarter: $turnover_{i,q} = 1/D \cdot \sum_{d=1}^{D} volume_{i,d}/(total\ \#\ of\ shares)_{i,d}$,	We take the trading volume and the total number of shares from Bloomberg. (Bloomberg Datatypes: PX_VOLUME and EQY_SH_OUT, respectively)	78,623

	where D is the total number of trading days during the quarter q.		
ILLIQ (Amihud,2002) $\ln(ILLIQ)_{i,q}$	The natural logarithm of the ILLIQ measure. ILLIQ of stock i for quarter q is the average of the daily ratios of the absolute level of the stock price change to the dollar volume, multiplied by a scaling factor of 10^6: $ILLIQ_{i,q} = 1/D \cdot \sum_{d=1}^{D} \left\|r_{i,d}\right\| / \$volume_{i,d} \cdot 10^6$, where D is the total number of trading days during the quarter q.	Stock prices from Bloomberg. (Bloomberg Datatype: PX_LAST) Share volumes from Bloomberg. (Bloomberg Datatype: PX_VOLUME)	79,719
Idiosyncratic Volatility $\ln(idio\_vol)_{i,q}$	The natural logarithm of idiosyncratic quarterly volatility of daily stock returns for each quarter. Idiosyncratic volatility is the standard deviation of daily risk-adjusted returns, estimated as the residuals of time-series regressions (over the whole sample) of the daily excess stock returns (over the risk-free rate) on the daily 4 factors of the Carhart model.	Stock prices are from Bloomberg. (Bloomberg Datatype: PX_LAST) The 4 factors (marker excess return, SMB, HML and MOM) and the risk-free rate come from the site of Kenneth French: http://mba.tuck.dartmouth.edu/pages/faculty/ken.french/data_library.html#Research	79,698
Excess market Return Rm_{q+1}-Rf_q	The excess market return is the value-weight return of all CRSP stocks that are incorporated in the US and are listed on NYSE, AMEX or NASDAQ and have share code 10 or 11 minus the risk-free rate (Treasury bill rate) for the relevant period.	Rm-Rf directly from the site of Kenneth French: http://mba.tuck.dartmouth.edu/pages/faculty/ken.french/data_library.html#Research	255 (monthly)
Small-minus-Big factor SMB_q	SMB is the return of a portfolio with long positions in small stocks and short positions in big stocks. The size break point is the median NYSE market equity.	SMB data directly from the site of Kenneth French: http://mba.tuck.dartmouth.edu/pages/faculty/ken.french/data_library.html#Research	255 (monthly)
High-minus-Low factor HML_q	HML is the return of a portfolio with long positions in value stocks and short positions in growth stocks. The book-to-market break points are the 30th and the 70th NYSE percentiles (below the 30th percentile are defined as the growth stocks and above 70th percentile are defined as the value stocks).	HML data directly from the site of Kenneth French: http://mba.tuck.dartmouth.edu/pages/faculty/ken.french/data_library.html#Research	255 (monthly)
Momentum factor MOM_q	MOM is the return of a portfolio with long positions in stocks with high prior returns and short positions in stocks with low prior returns. The monthly prior (2-12) return breakpoints are the 30th and 70th NYSE percentiles (below the 30th percentile are defined as the low prior return stocks and above 70th percentile are defined as the high prior return stocks).	MOM data directly from the site of Kenneth French: http://mba.tuck.dartmouth.edu/pages/faculty/ken.french/data_library.html#Research	255

Risk-free rate Rf_q	As Risk-free rate we use the one month Treasury bill rate.	Risk-free rate data directly from the site of Kenneth French: http://mba.tuck.dartmouth.edu/pages/faculty/ken.french/data_library.html#Research French takes the Treasury bill rate from Ibbotson Associates.	255 (monthly)
market beta / SMB beta / HML beta / MOM beta	Betas from rolling time-series regressions (with a 24-month window) of the monthly excess stock returns on the following four factors: Excess market return (Rm-Rf), SMB (Small-minus-Big), HML (High-minus-Low) and MOM (winner-minus-losers): $r_{i,m} - r_m^f = a + b_i^m(R_m - r^f)_m + b_i^{smb}(SMB)_m + b_i^{hml}(HML)_m + b_i^{mom}(MOM)_m + e_{i,m}$. We measure the monthly excess stock returns by subtracting from the monthly stock price changes the risk-free rate. We use the betas of the last month of each quarter to our analysis.	We take the Rm-Rf, SMB, HML, MOM and Rf data directly from the site of Kenneth French: http://mba.tuck.dartmouth.edu/pages/faculty/ken.french/data_library.html#Research Stock prices from Bloomberg. (Bloomberg Datatype: PX_LAST)	77,292 of each of the betas
Illiquidity $beta_q$	Illiquidity beta from rolling time-series regressions (with a 24-month window) of the monthly excess stock returns on the innovations of market-ILLIQ. In the same regression we also include Rm-Rf as an additional factor to control for the market comovement: $r_{i,m} - r_m^f = a + b_i^{illiq}(innov - mILLIQ)_m + b_i^m(R_m - r^f)_m + e_{i,m}$. The $mILLIQ$ is the cross-sectional mean of the $ILLIQ$, for each quarter q. The innovations of $mILLIQ$ are the residuals of an AR(1) model: $(mILLIQ)_m = c + (mILLIQ)_{m-1} + (innov - mILLIQ)_m$.	Rm-Rf and Rf data directly from the site of Kenneth French: http://mba.tuck.dartmouth.edu/pages/faculty/ken.french/data_library.html#Research Stock prices from Bloomberg. (Bloomberg Datatype: PX_LAST) Share volumes from Bloomberg. (Bloomberg Datatype: PX_VOLUME)	78,209

5. Descriptive Statistics

Figure 1 illustrates the stock ownership evolution of institutional investors over the sample period 1997-2016. Their participation increased from around 45% in 1997, to around 60% in 2000, reached 82% in 2012 and then stabilized around 78% after 2013. The whole distribution of institutional ownership keeps shifting to higher levels of participation from the first quarters of the sample to the later ones. The yellow boxes show that the middle 50% of the cross-sectional distribution was ranging between participations of 30% and 70% during the beginning of our sample, but afterwards it steadily shifted and after 2007 it is ranging between 65% and 95%. During the last years of the sample, the upper 25% of the distribution contains participations of above 95%. Notice also that the median of the cross sectional distribution is consistently above the mean and their gap goes up when the mean participation level rises after year 2000. These stylized facts are in line with the findings of earlier papers, which show the participation of institutional investors increases through time.

Figure 2 shows that institutional ownership is essentially divided up across 11 different styles, each with an average participation rate above 1%. The remaining 21 styles are small in size, having average participation rates of less than 1%. The biggest style is "Core Growth" with an average participation that exceeds 20%. Next to Core Growth is the "Index" style with average participation 18.7%, and is followed by "GARP" (18.3%), "Core Value" (14.9%), "Hedge Fund" (7.8%), "Deep Value" (7.4%), etc.

Figure 3 presents the distribution of the concentration parameter H of the different investment styles in a given stock in a given quarter. The distribution is over the pooled time series – cross sectional sample of 72,880 observations. Figure 3 shows a satisfactory dispersion of H across the pooled sample, enabling us to proceed with a meaningful econometric analysis. For the bulk of the stocks, H takes values between 0.12 and 0.35, a relatively wide range. As expected, the distribution of H is far from normal, yet it has a very long tail to the right. Later in the Appendix, we check the sensitivity of our econometric results to the presence of outliers in our main independent variable H.

Figure 4 traces the cross-sectional distribution of H over time. Mean concentration was gradually reduced from around 0.29 in the early years to slightly above 0.21 today. This is a substantial reduction in market-wide concentration, indicating that over the years, stocks are chosen by a more diversified pool of managers. The whole distribution of H shifts to

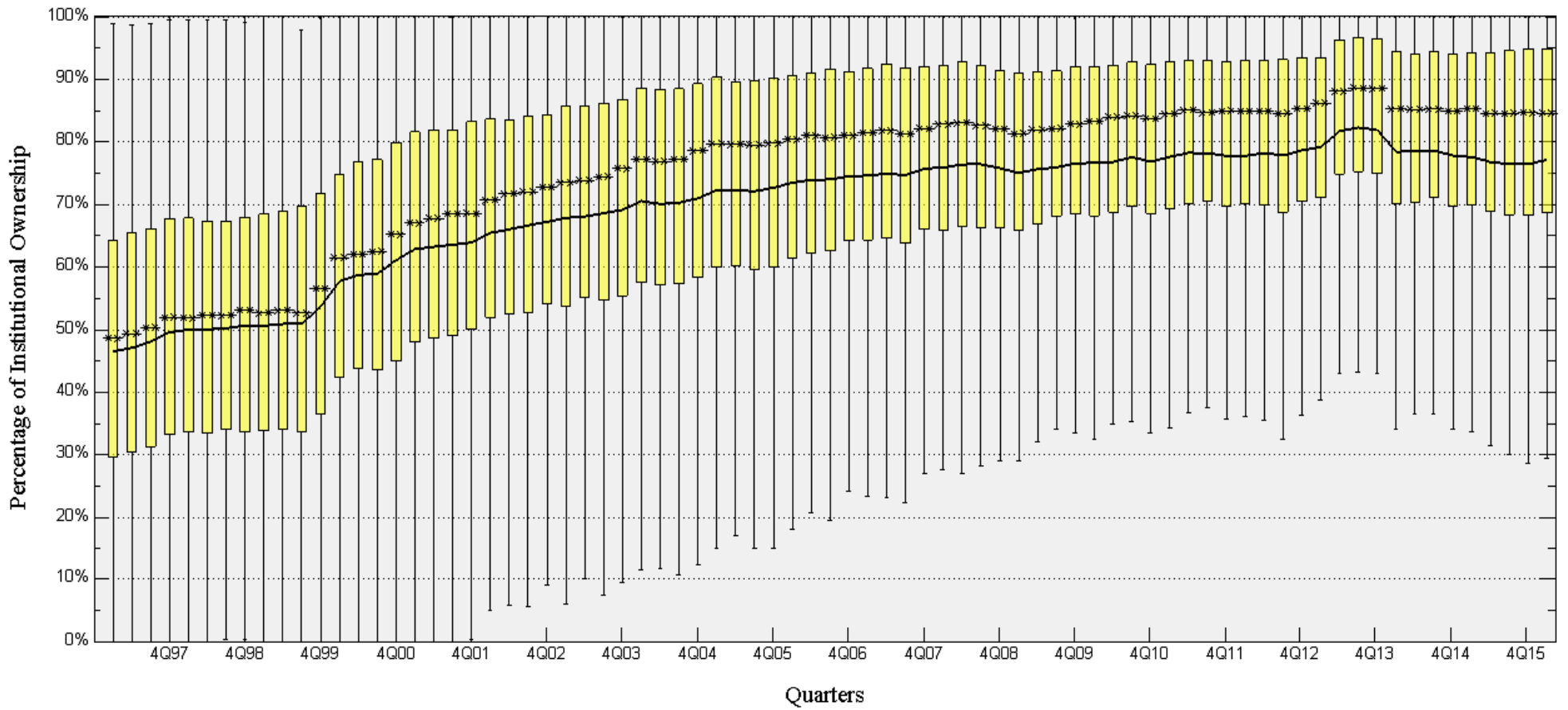
slightly lower levels and the range of the middle 50% of the distribution (yellow boxes) becomes narrower in the last quarters of the sample compared to the first quarters. These changes to the distribution of *H* are rather small and their overall effect on the econometric analysis limited.

Table 2 reports descriptive statistics of the main variables of our analysis and Table 3 does the same for the ownership shares of each of the 32 investment styles. Note that even the styles with very low average share of ownership, sometimes own a large nember of shares in at least some stocks. Hence, the maximum ownership can easily reach high values (last column).

Table 4 provides interesting evidence on the bivariate correlations of our independent variables. The correlation matrix has the concentration parameter *H* at the top. With minor exceptions, *H* is not highly correlated with the remaining independent variables. The most notable correlation of *H* is with ln(mv), the logarithm of market capitalization, and is -0.29. This negative correlation is expected, since bigger stocks are much more likely for be known and held by funds that follow distinctly different investment styles between them. *H* is also highly correlated with ln(ILLIQ). The correlation is positive at 0.43. To a large extent, this is a mechanical correlation, since by construction ILLIQ is highly correlated with size. Indeed, as shown in Table 4, the correlation between ln(mv) and ln(ILLIQ) is -0.87.

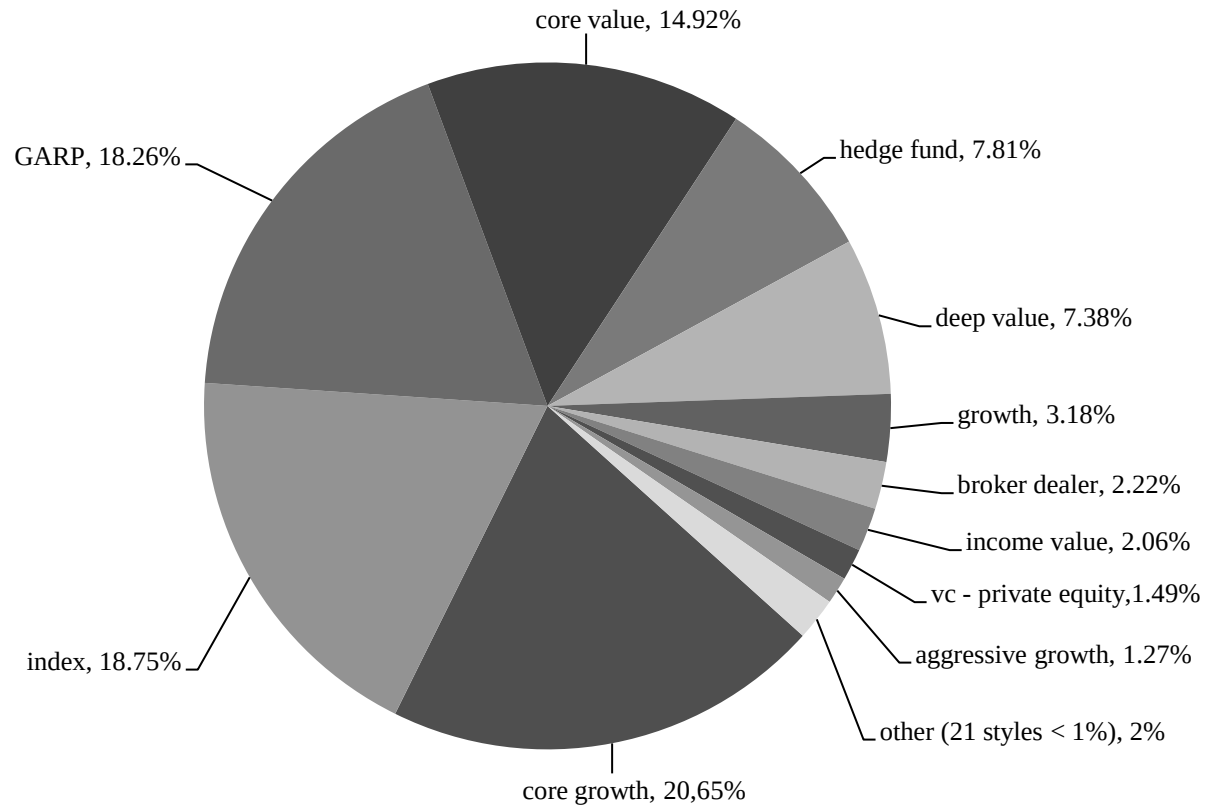
Table 5 contains the correlations of *H* with the stock ownership percentages of the large investment styles. As shown, *H* is not significantly correlated with any individual investment style. Its highest correlation is with the ownership of the Index style. This correlation is negative, at - 0.23. Apparently, a stock that is included in an index is widely known and thus it is more likely to be held by funds that follow distinctly different investment styles.

Figure 1: The evolution over time of the distribution of institutional stock ownership



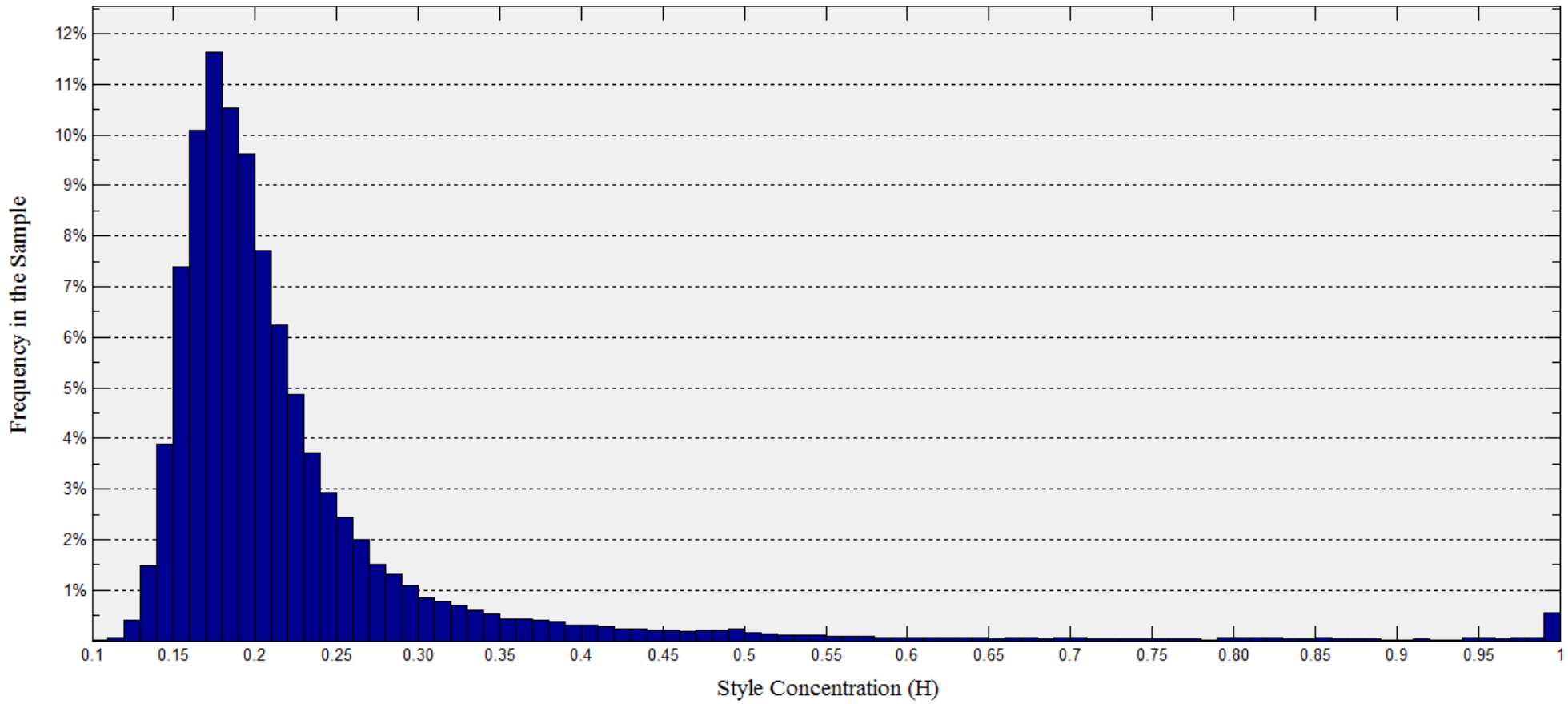
The figure illustrates the evolution of the distribution of institutional ownership over the 77 quarters of the sample (1997-Q1 to 2016-Q1). The solid black line represents the cross-sectional mean of institutional ownership for each quarter. Black stars represent the median institutional ownership in each quarter. The yellow boxes represent the middle 50% of the cross-sectional distribution of institutional ownership (from 25th percentile to 75th percentile). The black vertical lines above and below each yellow box cover a region of +/- 2.7 standard deviations above and below the mean of the cross-sectional distribution for each quarter.

Figure 2: Mean share of institutional ownership by investment style.



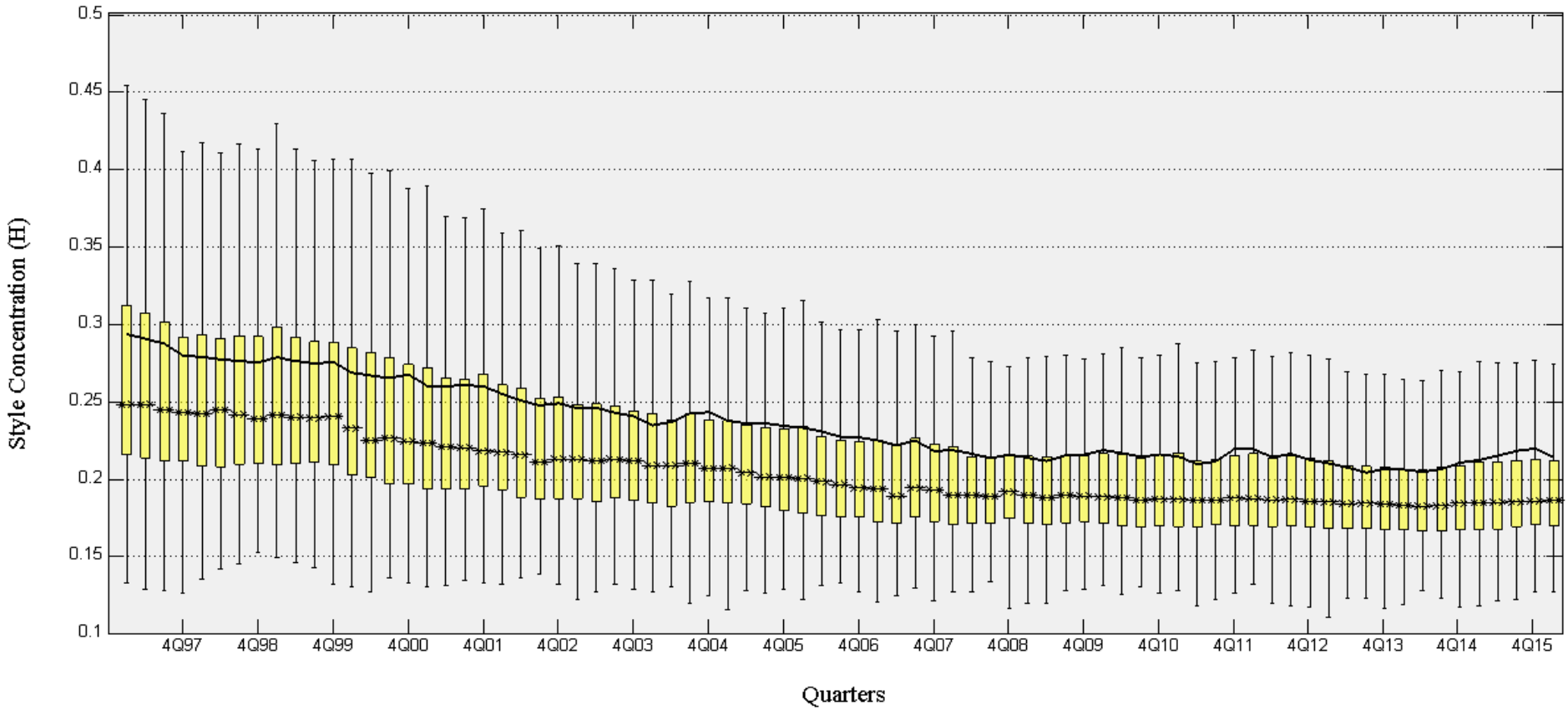
The figure illustrates the mean percentage shares of the investment styles in the pooled sample. Percentages add to 100%. . The sample covers 77 quarters, from the 1997-Q1 to the 2016-Q1 and contains 72,880 observations of stocks (an average of 946 stocks per quarter).

Figure 3: Distribution of the concentration (*H*) of investment styles in stock ownership



The figure illustrates the distribution of variable *H*, the concentration of investment styles in the ownership of stocks in the pooled sample. The sample covers 77 quarters, from the 1997-Q1 to the 2016-Q1 and contains 72,880 stock-quarters (on average 946 stocks per quarter). See Table 1 for the exact definition of *H*. The width of each bin is 0.01, thus in the figure there are 90 different bins from 0.10 to 1.00. The minimum value of H in the sample is 0.11 and the maximum is 1.

Figure 4:Evolution over time of the cross-sectional distribution of concentration parameter *H* in the institutional investment styles of stock ownership



The figure traces the evolution over time (from 1997-Q1 to 2016-Q1) of the cross sectional distribution of concentration parameter *H* in the institutional investment styles of stock ownership. The solid black line represents the cross-sectional mean of *H* in each quarter. Black stars represent the median *H* in each quarter. The yellow boxes represent the middle 50% of the cross-sectional distribution (from 25th percentile to 75th percentile). The black vertical lines above and below each yellow box cover a range of +/- 2.7 standard deviations above and below the mean of the cross-sectional distribution in each quarter.

Table 2: Descriptive statistics of main variables.

The table provides descriptive statistics over the pooled sample. The mean, standard deviation, skewness, kurtosis, minimum, median and maximum values are reported per variable. The definitions of the variables are described in Table 1.

	mean	*s.d.*	*skewness*	*kurtosis*	*min*	*median*	*max*
quarterly returns (%)	3.954	23.219	3.236	58.118	-97.976	3.061	821.875
H	0.233	0.118	3.949	21.774	0.111	0.199	1.000
mv ($bn.)	9.497	27.883	7.150	72.132	~0.000	1.965	572.283
ln(mv)	21.45	1.72	0.08	3.39	8.95	21.40	27.07
ln(mtb)	0.767	0.763	0.738	7.416	-6.725	0.704	8.379
ln(idiosyncratic volatility)	-4.038	0.538	0.294	3.687	-8.111	-4.063	-1.001
market beta	1.039	0.866	0.206	43.423	-25.534	0.989	27.909
SMB beta	0.527	1.230	1.509	23.916	-11.307	0.394	28.208
HML beta	0.356	1.341	-1.652	94.275	-65.941	0.312	15.470
MOM beta	-0.088	0.913	-0.418	12.887	-11.673	-0.043	12.601
standardized ln(ILLIQ)	~0.000	0.999	0.726	3.919	-2.529	-0.108	6.544
turnover (%)	0.791	1.893	86.286	9604.244	0.000	0.580	243.920
illiquidity beta	-0.113	1.683	-0.915	44.907	-58.811	-0.070	31.884
momentum	0.113	0.502	19.841	1677.949	-0.993	0.068	53.000
ln(dta)	-2.567	3.675	-3.136	11.695	-16.118	-1.434	1.559

Table 3: Descriptive statistics of the ownership percentages of each investment style

The table provides descriptive statistics of the ownership percentages of each investment style over the pooled sample. The mean, standard deviation, skewness, kurtosis, minimum, median and maximum values are reported per style. Percentages are based on the grand-total of shares of the 32 investment styles in each stock.

	mean	*s.d.*	*skewness*	*kurtosis*	*min*	*median*	*max*
Core Growth	20.65%	11.38%	1.81	10.35	0.00%	19.20%	100.00%
Index	18.75%	10.06%	1.18	9.87	0.00%	18.84%	100.00%
GARP	18.26%	11.38%	1.31	7.59	0.00%	16.88%	100.00%
Core Value	14.92%	10.68%	2.08	12.43	0.00%	13.19%	100.00%
Hedge Fund	7.81%	11.32%	3.61	21.30	0.00%	3.94%	100.00%
Deep Value	7.38%	7.86%	2.50	15.78	0.00%	4.97%	100.00%
Growth	3.18%	6.10%	8.07	99.27	0.00%	1.47%	100.00%
Broker – Dealer	2.22%	3.91%	8.71	127.31	0.00%	1.40%	100.00%
Income Value	2.06%	4.11%	9.12	154.14	0.00%	0.94%	100.00%
VC Private Equity	1.49%	8.26%	6.97	57.03	0.00%	0.00%	100.00%
Aggressive Growth	1.27%	3.04%	9.55	186.09	0.00%	0.27%	100.00%
Yield	0.84%	3.97%	14.54	255.07	0.00%	0.14%	91.50%
Specialty	0.66%	4.63%	14.92	263.71	0.00%	0.06%	100.00%
Momentum	0.18%	1.01%	13.14	329.09	0.00%	0.00%	49.94%
Sector Specific	0.12%	0.85%	18.33	510.87	0.00%	0.00%	42.79%
Long – Short	0.08%	0.96%	25.72	848.75	0.00%	0.00%	47.09%
Arbitrage	0.04%	0.28%	35.34	1,924.05	0.00%	0.00%	18.72%
Convertible Arbitrage	0.03%	0.50%	45.54	2,468.39	0.00%	0.00%	37.58%
Equity Hedge	0.02%	0.28%	41.17	2,362.67	0.00%	0.00%	26.00%
Event Driven	0.01%	0.53%	112.55	16,334.01	0.00%	0.00%	83.33%
Fixed Income Arbitrage	0.01%	0.37%	57.52	4,322.61	0.00%	0.00%	36.84%
Market Neutral	~0.00%	0.02%	12.63	239.99	0.00%	0.00%	0.78%
Emerging Markets	~0.00%	0.15%	52.97	3,099.51	0.00%	0.00%	11.36%

Table 3 (continued): Descriptive statistics of the ownership percentages of each style.

	mean	*s.d.*	*skewness*	*kurtosis*	*min*	*median*	*max*
Global Macro	~0.00%	0.04%	36.02	2,291.51	0.00%	0.00%	3.89%
Multi Strategy	~0.00%	0.38%	246.96	64,161.29	0.00%	0.00%	100.00%
Distressed	~0.00%	0.20%	158.58	29,645.82	0.00%	0.00%	41.21%
Funds of Funds	~0.00%	0.09%	187.13	35,474.75	0.00%	0.00%	16.33%
Mixed	~0.00%	~0.00%	19.14	451.67	0.00%	0.00%	0.13%
Emerging Market-Hedg.	~0.00%	0.02%	107.11	14,588.54	0.00%	0.00%	2.70%
CTA – Managed Fut.	~0.00%	~0.00%	75.93	7,060.72	0.00%	0.00%	0.23%
Quantitative	~0.00%	~0.00%	138.47	22,807.56	0.00%	0.00%	0.14%
Capital Struct. Arbitrage	~0.00%	~0.00%	113.23	13,050.04	0.00%	0.00%	0.07%

Table 4: Correlation matrix between the independent variables

Correlation Matrix between the independent variables, which are used in the econometric analysis. The sample covers 77 quarters, from the 1Q1997 to the 1Q2016 and includes on average 838 stocks per quarter.

	H	ln(mv)	ln(mtb)	ln (idio_vol)	market beta	SMB beta	HML beta	MOM beta	ln (ILLIQ)	turnover	illiquidity beta	momen-tum	ln (dta)
H	1												
ln(mv)	-0.293	1											
ln(mtb)	-0.058	-0.364	1										
ln(idio_vol)	0.117	-0.463	-0.143	1									
market beta	-0.065	-0.012	-0.065	0.156	1								
SMB beta	0.026	-0.287	-0.067	0.199	-0.134	1							
HML beta	0.019	-0.122	-0.148	0.093	0.080	0.052	1						
MOM beta	-0.011	0.072	0.090	-0.112	0.072	-0.132	0.080	1					
ln(ILLIQ)	0.430	-0.868	-0.309	0.333	-0.028	0.253	0.098	-0.060	1				
turnover	-0.075	-0.005	-0.065	0.127	0.062	0.037	-0.000	-0.025	-0.100	1			
illiquidity beta	0.002	0.026	0.011	-0.062	-0.016	-0.015	-0.017	-0.004	-0.018	-0.020	1		
momentum	0.001	0.041	0.218	-0.046	0.014	0.025	0.004	0.033	-0.038	0.007	-0.016	1	
ln(dta)	-0.030	0.153	-0.082	-0.098	0.006	-0.081	0.050	-0.025	-0.138	-0.005	0.000	-0.033	1

Table 5: Correlation matrix between the Style Concentration H and the shares of ownership of the ten biggest investment styles.

Correlation matrix between style concentration H and the percentage of holdings of the ten biggest investment styles. The sample covers 77 quarters, from the 1997-Q1 to the 2016-Q1 and on average includes 946 stocks per quarter.

	H	core growth	index	garp	core value	hedge fund	deep value	growth	broker-dealer	income value	VC – priv.equ.	aggr. growth
H	1											
core growth	0.076	1										
index	-0.227	-0.068	1									
garp	-0.081	-0.094	-0.176	1								
core value	-0.044	-0.173	-0.096	-0.184	1							
hedge fund	0.155	-0.293	-0.237	-0.230	-0.173	1						
deep value	-0.150	-0.110	-0.029	-0.173	-0.016	-0.097	1					
growth	0.053	-0.068	-0.149	0.015	-0.148	-0.058	-0.137	1				
broker-dealer	-0.066	-0.144	0.038	-0.142	-0.086	0.093	-0.086	-0.056	1			
income value	-0.056	-0.045	0.069	-0.094	-0.048	-0.107	-0.018	-0.067	-0.032	1		
VC – private equity	0.200	-0.180	-0.214	-0.136	-0.136	0.021	-0.114	-0.019	0.005	-0.063	1	
aggressive growth	-0.063	-0.016	-0.100	0.040	-0.092	-0.019	-0.119	0.063	-0.039	-0.057	-0.004	1

6. Econometric analysis

6.1 Equation specification and control variables

We now investigate the relation between style concentration H and stock returns. The nature of this relation is predictive, thus the basic test is between the concentration of stock i at the end of the quarter q, $H_{i,q}$, and the quarterly stock returns of stock i during quarter q+1, $r_{i,q+1}$:

$$r_{i,q+1} = \alpha + \beta \cdot H_{i,q} + e_{i,q+1} \qquad (3)$$

The empirical hypothesis, which is based on Merton's prediction, is that higher concentration predicts higher expected returns, hence:

$$Hypothesis\colon\ \beta > 0 \qquad (4)$$

We furthermore use a host of control variables, which are either directly linked with Merton's model, or are known characteristics related to asset pricing anomalies, or are related to specific styles:

$$r_{i,q+1} = \alpha + \beta \cdot (H)_{i,q} + \Gamma' \cdot Z_{i,q} + e_{i,q+1,} \qquad (5)$$

where Γ is a vector of coefficients for the control variables, and Z is a matrix that contains the control variables. All the controls variables are measured during quarter q.

We include as a first control variable the market beta, the traditional milestone risk factor in asset pricing models (CAPM of (Sharpe (1964), Lintner (1965) and Mossin (1966)). Market beta is also included in the analysis of Merton (1987). Merton's model simplifies to the traditional standard CAPM if information were complete and all investors have full information about all the existing stocks. As further controls we also include the three additional betas of the Carhart (1997) four-factor model, i.e., the SMB beta, the HML beta and the MOM beta. This four-factor model captures the exposure of a stock to systemic risk more fully. Also, these additional three betas can be thought to be proxies of certain investment style returns, as they are constructed as zero cost portfolios, sorted on the same characteristics that define the styles (i.e. the size, the market-to-book ratio and the momentum). By including them in the regresion, we have the extra benefit of also controlling for possible side effects of specific styles.

Next, we also include the logarithm of the market capitalization as a control variable in our specification. Previous empirical studies find that market capitalization is significantly and negatively correlated with future returns and the size anomaly is still present today.[25] Size is also one of the key variables of Merton's model. According to his model the size should have positive correlation with future stock returns. However, as Merton discusses, in reality the size is correlated with a number of other variables, including the concentration, the volatility and the illiquidity of a stock. He goes at length to explain that even if the relation $\partial r/\partial(size) > 0$ holds, the $dr/d(size)$ could be negative.[26] Finally, the stock size is an important characteristic for the quantitative determination of the styles (Brown and Goetzman (1997), Chan et al. (2002), Barberis and Shleifer (2003), Teo and Woo (2004) and Wahal and Yavuz (2013), among others). Hence, controlling for it (in addition to the control for the effect of the beta of SMB risk factor) provides additional confidence that our results are not driven by the size effect or by any specific style, which is defined along this characteristic.

We include the logarithm of the idiosyncratic volatility of returns as an additional control variable. Merton (1987) provides the theoretical underpinnings for the relation between idiosyncratic volatility and expected returns. In an economy in which investors do not hold fully diversified portfolios, the idiosyncratic price volatility should have a positive relation with expected returns in order to reward investors for the excess amount of risk they undertake by being away from their optimal portfolios. In the empirical literature, the debate about the role of idiosyncratic volatility remains open. Lintner (1965), Lehmann (1990), Tinic and West (1986), Melkiel and Xu (2002) and Fu(2009) find that the relation between volatility and stock returns is positive. However, Shleifer and Vishny (1997) argue that arbitrageurs do not trade stocks with higher idiosyncratic volatility, due to the higher probability for these stocks to move further away from fundamentals before they converge back to them, and thus they remain overvalued. As a result, these stocks exhibit lower future returns. Ang et al. (2006) confirm the hypothesis of Shleifer and Vishny. Following the

[25] The size effect is present in a very large number of papers. The first papers that formally indicate the existence of the relationship between size and stock returns was that of Basu (1977) and Banz (1981). Jegadeesh (1990), Fama and French (1992) and Brennan et al. (1998) also find that the size effect is significantly and negatively correlated with the stock returns. Avramov and Chordia (2006) in a more recent paper still find that size effect is significant.

[26] See Merton (1987) p.497.

original Merton (1987) model, we include the variable as an additional contol variable in our analysis.

The logarithm of the ratio of market-to-book value is also included as a control variable. The literature finds that the book-to-marker ratio (the reciprocal of the ratio that we use) is significantly and positively correlated with expected returns.[27] It is important to control for this variable because, in addition to size, the market-to-book ratio is another stock characteristic that influences the choice of investment style (see Brown and Goetzman (1997), Chan et al. (2002), Barberis and Shleifer (2003), Teo and Woo (2004) and Wahal and Yavuz (2013), among others).

As further controls, we use two measures of stock illiquidity, first, ILLIQ, which is the priced impact measure of Amihud (2002) and second, the share turnover. The positive relation between illiquidity and stock returns is well documented by the relevant literature (Amihud and Mendelson (1986), Brennan and Subrahmanyam (1996), Brennan et al. (2012) among others). And the high correlation of our main independent variable, the style concentration parameter *H*, with ILLIIQ, which was documented earlier in Table 4, makes it imperative to include ILLIQ as a control variable in the econometric analysis. The second liquidity measure which we use as a control variable is turnover. There is a well documented strong negative relation between share turnover and stock returns (see Brennan et al. (1998), Avramov and Chordia (2006)). Finally, in order to capture the effects of illiquidity fully, we include a third variable, the illiquidity beta as a measure of illiquidity risk. Amihud (2002), Pastor and Stambaugh (2003) and Acharya and Pedersen (2005) and find that stocks with higher illiquidity risk have higher expected returns.

In the analysis we also include the price momentum as an extra control variable. Price momentum is a variable that is positively correlated with future stock returns (Jegadeesh and Titman (1993), Brennan et al. (1998), Avramov and Chordia (2006), among others). Moreover, momentum is also related with the institutional trading and the style investing (Grinblatt et al. (1995), Wermers (1999), Nofsinger and Sias(1999), Badrinath and Wahal (2002) and Chan et al. (2002), among others). We thus control for the momentum in order to ensure that our results are not driven by any momentum effect.

[27] See Ball (1978), Fama and French (1992), Brennan et al. (1998), Avramov and Chordia (2006), Fama and French (2015), among others.

We use as an extra control variable the debt-to-asset ratio (leverage). Bhandari (1988) finds that a measure of leverage is positively related with the future stock returns. Fama and French (1992) also find evidence about a relation between leverage and stock returns, yet they also find that – to a large degree – the size and book-to-market variables absorb the effect of leverage.[28]

Finally, we pay close attention to style effects, which could co-exist with firm effects and confound the influence of ownership concentration *H*. For example, if a stock were followed by a specific style and the returns of that style were exhibiting momentum, it is possible the returns of the stock would be positively affected, even if the stock itself has no momentum at the individual level. Net style inflows are also are positively correlated with future stock returns.[29] To avoid the confounding, we use the ownership percentages of each of the 32 investment styles that may be present in each stock as additional control variables. We are thus in position to directly control for any effect associated with a specific style, which is not already captured by the previously mentioned firm-related characteristics.

6.2 Main Econometric Results

We run pooled time series – cross sectional OLS regressions, including 75 quarterly dummies in order to address the time effect.[30] As a consequence of the time effect, the observed Adjusted R^2s are unusually high. In the quarterly horizon of Table 6, they range from 21% to 24%. In other words, the high explanatory power of the regressions is primarily due to a common shift from quarter to quarter of the dependent variable and the independent variables. Observe that the t-statistics in parentheses below the coefficient estimates are based on White heteroskedasticity-consistent standard errors. The standard White (1980) correction addresses the heteroskedasticity, which is present in our data, and corrects (reduces) the size of the reported t-statistics. For easiness of exposition, we use three

[28] Besides, there is evidence that higher leverage in value firms has negative effect to their future price. Piotroski (2000) use the leverage (along with other firm specific characteristics) to measure the financial soundness of a firm.

[29] Barberis and Shleifer (2003), Teo and Woo (2004) and Froot and Teo (2008) provide theoretical underpinnings and empirical evidence of these style effects.

[30] There are 76 quarters available for estimation, one quarter less than the available data on concentration parameter *H*.

asterisks (***) to denote statistical significance at the 1% level, two asterisks (**) at the 5% level, and one asterisk at the 10% level.

The quarterly stock returns are not serially correlated. Hence, there is no need to address the possibility of a firm effect (Cochrane (2001), Petersen (2009)). We report annualized parameter estimates, so the parameters are more easily interpretable and comparable with other results in the literature. Overall, the estimation results in Table 6 provide overwhelming support for the hypothesis that style concentration of ownership is positively correlated with future expected stock returns, a hypothesis consistent with the prediction of Merton (1987). The coefficient estimates of the "influence" of style concentration *H* on expected stock returns are both economically and statistically significant.

Table 6 includes ten sets of parameter estimates from ten different regressions for the quarterly horizon. In the first regression (in column 1), the only independent variable (in addition to the constant term and the 75 quarterly dummy variables) is the style concentration parameter $H_{i,q}$. In the 10^{th} regression (in column 10), $H_{i,q}$ is accompanied by the full set of control variables. The in-between columns (2 to 9) provide information on various interesting combinations of the independent variables.

Column 1 of Table 6 shows the regression coefficient β of style concentration parameter $H_{i,q}$ to be 21.9 with a t-statistic of 5.26. Given the non-linear nature of *H*, the interpretation of β requires care. When our independent variable *H* moves drastically from its minimum value of 0.11 to its maximum value of 1.00 within quarter q, then next quarter's annualized return at q+1 is expected to increase on average by 19.5 percentage points (21.9 X(1.00-0.11)), which is huge. For more realistic changes in *H*, say, a one standard deviation increase in *H* of 0.12 units, the average increase in expected returns is 2.63%. This is an economically significant change.

Column 2 of Table 6 adds to the previous regression in column 1, the 32 percentage ownership shares of each investment style in each stock-quarter. We thus test whether the effect of style concentration *H* reflects true inattention in stocks or, instead, is affected by the influence of various investment styles. It is reassuring that the new regression coefficient β does not change much and remains high at 18.5 with a t-statistic of 4.07. The result ensures that it is the concentration (or participation in the words of Merton) rather than any specific investment style that drives the correlation of $H_{i,q}$ with $r_{i,q+1}$.

Column 3 of Table 6 adds to the original regression in column 1, the CAPM beta. It is significant only at the 10% level, while the coefficient β of $H_{i,q}$ remains approximately the same, at 23.3 with a t-statistic of 5.23. The lack of strength of the CAPM beta is consistent with earlier evidence on this issue.

Column 4 adds to the variables in column 3, three more beta factors, SMB (Small minus Big), HML (High minus Low), and MOM (momentum). In the benchmark four factor model of column 4, the coefficient β of $H_{i,q}$ remains approximately the same at 22.9 with a t-statistic 5.16. Note that Market beta enters the regression with a positive and statistically significant coefficient only at the 10% level. The SMB beta enters the regression with a positive but not statistically significant coefficient, the HML with a positive and significant one, while the MOM coefficient with a negative and significant coefficient.

Column 5 includes the four variables of the theoretical model of Merton: The participation variable, proxied by our style concentration *H*, the market beta, the size, ln(mv), and idiosyncratic volatility, ln(idio_vol). Column 5 can be compared with column 3, which only includes two of the four Merton variables. Note that regression coefficient β of style concentration *H* decreases in magnitude, at 9.81 with a t-statistic of 2.20. Apparently, size and volatility, being correlated with *H*, take away some of the explanatory power of style concentration. Recall from Table 4 that *H* has a negative correlation with size of -0.29 and a positive correlation with idiosyncratic volatility of 0.12. Size and idiosyncratic volatility are themselves negatively correlated at -0.46. Yet, as we see later, this is the lowest value that β takes across all of our ten regressions. When more controls are added to the regression, the size of β gets reinstated, especially in the full-blown model in column 10.

In column 5, the coefficient of the size variable is negative and highly statistically significant, a result which is in line with the findings of the size effect in the literature.[31] Although Merton's model predicts that the relative size of a company should positively predict subsequent stock returns, this is not the case in any of the empirical studies. Merton is aware of the problem and highlights that size could be simultaneously an inverse proxy for

[31] Small firms have higher future returns relative to large firms. The negative coefficient on the ln(mv) variable remains similar in magnitude in the full-blown model of column 10, but its t-statistic declines to about a third its original value in clolumn 5, i.e., to -4.53 from -11.80.

volatility (which should be positively correlated with stock returns according to his model) and thus even if $\partial r/\partial(size) > 0$ holds, the $dr/d(size)$ could be negative.[32]

In column 5, the coefficient of idiosyncratic volatility is positive and significant, confirming the prediction of Merton about the positive relation between idiosyncratic price volatility and stock returns. This result is in line with the empirical findings of Lintner (1965), Lehmann (1990), Tinic and West (1986), Melkiel and Xu (2002) and Fu(2009), who find that idiosyncratic volatility is positively correlated with future stock returns.[33] Finally, in this specification, the coefficient of market beta is positive but insignificant, confirming the findings of the literature about the empirical weakness of CAPM.

Column 6 adds to the previous Merton specification the market-to-book variable, ln(mtb). This variable is a major determinant of investment styles and by including it, one can control for a possible confounding influence, originating from style strategies. The addition of the market-to-book variable does not change much the earlier results. The coefficient of style concentration become a bit stronger and the coefficient of size is reduced, while the coefficients of volatility and market beta remain about the same.

Column 7 adds the three additional betas of the four-factor model of Carhart (1997). The style concentration coefficient does not change much.[34] Column 8 presents an alternative specification to column 7. Instead of adding factor betas, it adds to column 6, two liquidity variables, ln(ILLIQ) and share turnover. Now the coefficient β of style concentration rises substantially to 13.47.[35]

---

[32] Of course in our regression in column 5, we control for idiosyncratic volatility and size continues to be associated negatively with future returns.

[33] Our results are not in line with the result of Ang et al. (2006) who find that idiosyncratic volatility is negatively correlated with future stock returns.

[34] Yet one can also compare the specification in column 7 with the specification in column 4, which does not include the variables of size, idiosyncratic volatility, and market-to-book ratio. Interestingly, now the coefficient of the SMB beta is negative and significant (apparently due to the simultaneous presence of size) and the coefficient of HML beta is now insignificant (apparently due to the presence of market-to-book ratio).

[35] Notice that turnover enters the regression with a negative and significant coefficient, confirming the findings of the relative literature. However, the coefficient of ILLIQ is negative (but with t-statistic -1.69), a result opposite to that of the existing literature. The explanation for this result is the simultaneous existence in the regression of the variables of size, volatility and turnover, which are basically the determinants of ILLIQ. The negative coefficient of the remaining ILLIQ effect is possibly due to very illiquid stocks, which converge very slowly to their fundamental value and thus they appear with a persistent undervaluation. When we include to the regression the ln(ILLIQ) without the log(mv), the coefficient of the former is positive and significant, which is in line with the empirical evidence of the relevant literature.

Column 9 presents the regression with all the control variables present, except for the ownership percentage share of each style. Column 9 also adds three more controls we have not encountered thus far: illiquidity beta, price momentum and the ln(dta)). Notice that the coefficient β rises even more relative to column 8. It is now 15.15 with a t-statistic of 2.48.

Column 10 presents the results for the full specification. It adds to column 9 the 32 variables of the shares of the investment styles we had seen earlier in column 2. (In reality, we add 31 share variables to avoid complete multi-collinearity). This last regression controls directly for any possible confounding influence on *H* originating from the investment strategies themselves. It turns out they have no effect on the estimates of regression coefficient β. If anything, the coefficient now gets a bit bigger, at 17.46, with a t-statistic pf 2.52. This size of β translates into an annual premium of almost 2.10% for a one standard deviation increase in *H*. This is a very large premium, especially when one considers the fact that it comes on top of the premia for a large number risk factors and other determinants of expected stock returns, as already shown in the regression of column 10.

Table 6: Stock returns and previous quarter's style concentration in ownership

Panel OLS regressions of the annualized quarterly stock i return at quarter q+1, $r_{i,q+1}$, on the style concentration in ownership of stock i, $H_{i,q}$, of the previous quarter q, and on other lagged control variables for stock i, $Z_{i,q}$, which are also observed during quarter q:

$$r_{i,q+1} = \alpha + \beta \cdot (H)_{i,q} + \Gamma' \cdot Z_{i,q} + e_{i,q+1}.$$

There are 10 regressions in columns 1 through 10. A time effect with quarterly dummies is included in every regression. The variables of each regression are described in the very left column. The variables denoted as "% Style Ownership" are the ownership percentage shares of 32 different investment styles (we include only 31 of the 32 styles to avoid perfect multicollinearity). The variables denoted as "Other Controls" are the following: illiquidity beta, price momentum and the ln(dta)). See Table 1 for the detailed definitions of the variables.

Returns are measured in percentage form. The sample covers the period between 1997-Q1 and 2015-Q4 (76 quarters) and on average consists of around 838 stocks in each quarter. The total number of observations in each regression is described in the last row. t-statistics are inside the parentheses below the regression coefficients, which are based on White (1980) heteroskedasticity-consistent standard errors. Three asterisks *** denote statistical significance at the 1% level, two asterisks ** at the 5% level, and a single asterisk * at the 10% level. Adj-R^2 is the adjusted coefficient of determination of the regression, expressed in %.

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
H	**21.92*** (5.26)**	**18.52*** (4.07)**	**23.27*** (5.23)**	**22.92*** (5.16)**	**9.81** (2.20)**	**10.56** (2.45)**	**9.97** (2.31)**	**13.47*** (2.59)**	**15.15** (2.48)**	**17.46** (2.52)**
ln(mtb)			-	-	-	-3.56*** (-5.31)	-3.31*** (-4.81)	-3.68*** (-5.48)	-3.14*** (-5.05)	-3.59*** (-5.38)
ln(idio_vol)			-	-	6.13*** (5.14)	6.13*** (5.38)	6.33*** (5.39)	6.52*** (6.59)	6.67*** (5.75)	5.34*** (4.42)
ln(mv)			-	-	-2.65*** (-11.80)	-1.94*** (-8.83)	-2.14*** (-8.97)	-2.65*** (-5.40)	-2.98*** (-5.28)	-2.63*** (-4.53)
market beta			1.14* (1.92)	1.18* (1.94)	0.19 (0.32)	0.26 (0.46)	0.06 (0.11)	0.28 (0.48)	-0.13 (-0.23)	-0.35 (-0.61)
SMB beta				0.49 (1.17)			-1.02** (-2.21)	-	-1.09** (-2.35)	-1.15** (-2.47)
HML beta				1.13*** (2.96)			0.25 (0.66)	-	0.38 (1.11)	0.45 (1.30)
MOM beta				-1.71*** (-3.17)			-1.08** (-2.17)	-	-1.07** (-2.11)	-1.04** (-2.06)
ln(ILLIQ)								-1.51* (-1.69)	-1.83* (-1.75)	-1.44 (-1.30)
turnover								-0.76*** (-4.39)	-0.75*** (-4.39)	-0.78*** (-4.35)
% Style Ownership	-	YES	-	-	-	-	-	-	-	YES
Other Controls	-	-	-	-		-	-	-	YES	YES
Adj-R^2 (%)	21.3	21.5	21.1	21.3	21.8	22.9	23.0	23.0	23.6	23.7
Number of observations	70,490	70,490	67,881	67,881	66,971	64,807	64,807	64,802	63,704	63,704

7. Multi-year horizons

We now repeat the earlier analysis by measuring stock returns over multi-period horizons and running the following regression:

$$r_{i,q+k} = \alpha + \beta \cdot (H)_{i,q} + \Gamma' \cdot Z_{i,q} + e_{i,q+k,} \qquad (6)$$

where k = 4, 8, 12, 16 quarters, the independent variables are the same as before and are observed in quarter q, and the multi-period return $r_{i,q+k}$ is the cumulative product of the individual gross quarterly returns, annualized and observed at the end of quarter q+k.

Our sample continues to be quarterly and is of the same approximate size as the sample of the quarterly horizon. Recall in the quarterly horizon of Table 6, we lost one quarter's worth of data in order to measure stock returns one quarter later, namely we lost the last quarter of the sample, 2016-Q1. In Table 6, the sample ended in 2015-Q4. Here, in the annual horizon, with k = 4, we lose 4 data points per stock, and the sample ends in 2015-Q1. Similarly, in the horizon of four years ahead, with k = 16, the longest we examine, we lose 16 data points per stock, and the sample ends in 2012-Q1. We use panel Newey-West (1987) standard errors, which correct both for conditional heteroskedasticity and for the serial correlation of the residuals. This serial correlation is not present in the quarterly horizon of Table 6, yet it is being introduced mechanically from the overlap of the periods over which we measure stock returns.

The aim of the multi-year analysis is to investigate whether the effects of style concentration on current prices and, hence, on future returns, are temporary or more durable. If the effect on prices were temporary due to style-related strategies, then prices would correct immediately and the effect on returns would disappear or even reverse as the horizon gets extended to periods longer than a quarter. If, however, the effect originates from Merton's lack of participation, then the effect on returns can last as long as the dislocation effect on equilibrium prices persists. If the dislocation remains the same or disappears slowly over time, then the effect on multiperiod annualized returns remains present, but becomes smaller in size as the horizon grows. If, however, the dislocation grows bigger for a period longer than a quarter, then the style concentration effect on future annualized multi-year returns can even grow in size.

We use annualized stock returns that are measured 1, 2, 3 and 4 years ahead as the dependent variables, and repeat at the quarterly frequency our earlier econometric analysis

for those returns. Table 7 presents the results for returns calculated over 1 and 2 years ahead, while Table 8 presents the equivalent results for 3 and 4 years ahead.

We repeat the most important five of the ten earlier regressions in Table 6, namely the regressions contained in columns 1, 2, 6, 9, and 10. The univariate regression is in column 1. In this regression, the coefficient β of the concentration variable $H_{i,q}$ remains positive and statistically significant in all four horizons of 1, 2, 3, and 4 years ahead. The size of β is of great interest, as it rises over time despite the fact that returns are annualized! In the annual horizon it equals 24.0, compared to 21.9 of the quarterly horizon. This implies that the effect one year later is four times larger than the effect next quarter, showing that the market adjusts slowly and in the same direction to the original shock in $H_{i,q}$. More surprisingly, the effect continues growing over years 2, 3 and 4. The two-year β is 29.0, the three-year 29.7, and the four-year 34.7. Put differently, the four-year effect is at least twice as big as the two-year effect, which in turn is at least twice as big as the one-year effect, which is at least four times as big as the quarterly effect!

Columns 2 in Tables 7 and 8 add to the univariate case the shares of the individual investment styles. Now the estimation controls for possible confounding effects on $H_{i,q}$ originating from the style strategies. We find the same result we found earlier in Table 6. Namely, the coefficient β of $H_{i,q}$ does not change much relative to the simple univariate case, and particularly in the longer horizons it remains practically the same as in the univariate case.

Columns 6 in Tables 7 and 8 present the variables of the Merton model, enhanced with the market-to-book variable, which is an important variable in the choice of investment style. The new β estimates of variable $H_{i,q}$ are weaker relative to the univariate case, yet stronger relative to the quarterly horizon of Table 6.

Columns 9 of Tables 7 and 8 contain the full model, with all the control variables except for the 32 style shares. Columns 10 include the style shares as well. Again, there are no surprises. The coefficient β of the variable of interest, $H_{i,q}$, remains statistically significant at the 5% level up to three years ahead, and at the 10% level in the four-year horizon.

Regarding the remaining variables in Tables 7 and 8, the coefficients of market-to-book ratio and turnover are reduced (and their t-statistics as well), the coefficients of volatility and size remain around the same level (both in terms of point estimate and of t-

statistics) and that of market beta turns to positive but it is still insignificant. Overall, the multi-year horizon results in Tables 7 and 8 provide strong support for the Merton model and the role of style concentration in stock ownership in explaining the cross-section of expected stock returns.

Table 7: Multi-year Stock Returns (1-year and 2-years ahead) and past style concentration

Panel OLS regressions of the annualized stock i return $r_{i,q+k}$ from the end of quarter q to the end of quarter q+k, where k = either 4 or 8, on the style concentration in ownership of stock i, $H_{i,q}$, of quarter q, and on other control variables for stock i, $Z_{i,q}$, which are also observed during quarter q:

$$r_{i,q+k} = \alpha + \beta \cdot (H)_{i,q} + \Gamma' \cdot Z_{i,q} + e_{i,q+k}.$$

There are 5 columns per horizon k, which correspond to the columns in Table 6. A time effect with quarterly dummies is included in every regression. The variables of each regression are described in the very left column. The variables denoted as "% Style Ownership" are the ownership percentage shares of 32 different investment styles (we include only 31 of the 32 styles to avoid perfect multicollinearity). The variables denoted as "Other Controls" are the following: illiquidity beta, price momentum and the ln(dta)). See Table 1 for the detailed definitions of the variables.

Returns are measured in percentage form. The sample covers the period from of 1997-Q1 to 2015-Q1 (for k = 4) or 2014-Q1 (for k = 8). The quarterly cross section consists on average of around 818 stocks in the one-year horizon, and 794 stocks in the two-year horizon. The total number of observations in each regression is described in the last row. t-statistics are inside the parentheses below the regression coefficients, which are based on Newey and West (1987). Three asterisks *** denote statistical significance at the 1% level, two asterisks ** at the 5% level, and a single asterisk * at the 10% level. Adj-R^2 is the adjusted coefficient of determination of the regression, expressed in %.

	(1)	(2)	(6)	(9)	(10)	(1)	(2)	(6)	(9)	(10)
	1y-Raw Returns	**1y-Raw Returns**	**1y-Raw Returns**	**1y-Raw Returns**	**1y-Raw Returns**	**2y-Raw Returns**	**2y-Raw Returns**	**2y-Raw Returns**	**2y-Raw Returns**	**2y-Raw Returns**
H	**23.99*** (6.14)**	**18.94*** (4.71)**	**15.18*** (3.41)**	**14.75*** (2.95)**	**18.66*** (2.96)**	**29.01*** (5.13)**	**25.23*** (4.37)**	**14.13*** (2.61)**	**15.13** (2.53)**	**19.79** (2.49)**
ln(mtb)			-2.85*** (-5.10)	-2.57*** (-4.53)	-1.98*** (-3.42)			-2.61*** (-4.14)	-2.42*** (-3.84)	-1.54** (-2.27)
ln(idio_vol)			4.61*** (4.81)	5.29*** (5.16)	4.87*** (4.57)			5.60*** (4.94)	6.06*** (4.91)	5.91*** (4.58)
ln(mv)			-2.23*** (-9.69)	-2.55*** (-5.09)	-2.67*** (-5.03)			-2.63*** (-8.18)	-3.15*** (-4.75)	-3.59*** (-5.50)
market beta			1.18** (2.15)	0.79 (1.42)	0.63 (1.12)			0.80 (1.21)	0.55 (0.83)	0.39 (0.56)
SMB beta				-1.29*** (-3.31)	-1.36*** (-3.33)				-0.90** (-2.08)	-1.07** (-2.31)
HML beta				0.83** (2.28)	0.85** (2.39)				0.54 (1.23)	0.56 (1.32)
MOM beta				-0.78* (-1.88)	-0.63 (-1.46)				-0.63 (-1.18)	-0.44 (-0.82)
ln(ILLIQ)				-0.27 (-0.32)	-0.63 (-0.64)				-0.78 (-0.79)	-1.72 (-1.54)
turnover				-0.43** (-2.11)	-0.42** (-2.09)				-0.38** (-2.35)	-0.38** (-2.24)
% Style Ownership	-	YES	-	-	YES	-	YES	-	-	YES
OtherControls	-	-	-	-	YES	-	-	-	-	YES
Adj-R^2(%)	14.0	14.5	19.5	19.6	20.0	10.4	11.1	15.1	15.2	15.7
Number of observations	65,589	65,589	59,919	59,914	58,889	60,466	60,466	54,955	54,949	53,960

Table 8: Multi-year Stock Returns (1-year and 2-years ahead) and past style concentration

Panel OLS regressions of the annualized stock i return $r_{i,q+k}$ from the end of quarter q to the end of quarter q+k, where k = either 12 or 16, on the style concentration in ownership of stock i, $H_{i,q}$, of quarter q, and on other control variables for stock i, $Z_{i,q}$, which are also observed during quarter q:

$$r_{i,q+k} = \alpha + \beta \cdot (H)_{i,q} + \Gamma' \cdot Z_{i,q} + e_{i,q+k}.$$

There are 5 columns per horizon k, which correspond to the columns in Table 6. A time effect with quarterly dummies is included in every regression. The variables of each regression are described in the very left column. The variables denoted as "% Style Ownership" are the ownership percentage shares of 32 different investment styles (we include only 31 of the 32 styles to avoid perfect multicollinearity). The variables denoted as "Other Controls" are the following: illiquidity beta, price momentum and the ln(dta)). See Table 1 for the detailed definitions of the variables.

Returns are measured in percentage form. The sample covers the period from of 1997-Q1 to 2013-Q1 (for k = 12) or 2012-Q1 (for k = 16). The quarterly cross section consists on average of around 772 stocks in the three-year horizon, and 756 stocks in the four-year horizon. The total number of observations in each regression is described in the last row. t-statistics are inside the parentheses below the regression coefficients, which are based on Newey and West (1987). Three asterisks *** denote statistical significance at the 1% level, two asterisks ** at the 5% level, and a single asterisk * at the 10% level. Adj-R^2 is the adjusted coefficient of determination of the regression, expressed in %.

	(1)	(2)	(6)	(9)	(10)	(1)	(2)	(6)	(9)	(10)
	3y-Raw Returns	**3y-Raw Returns**	**3y-Raw Returns**	**3y-Raw Returns**	**3y-Raw Returns**	**4y-Raw Returns**	**4y-Raw Returns**	**4y-Raw Returns**	**4y-Raw Returns**	**4y-Raw Returns**
H	**29.67*** (5.34)**	**28.12*** (4.64)**	**15.13*** (2.65)**	**15.17** (2.32)**	**18.37** (2.12)**	**34.70*** (4.98)**	**34.14*** (4.11)**	**20.93*** (2.61)**	**19.95** (2.05)**	**17.14* (1.74)**
ln(mtb)			-2.44*** (-4.10)	-2.10*** (-3.46)	-1.43** (-1.97)			-3.58*** (-5.15)	-3.13*** (-4.27)	-2.32*** (-2.83)
ln(ido_vol)			5.20*** (4.32)	5.53*** (4.38)	4.93*** (3.95)			7.07*** (4.96)	7.46*** (4.88)	6.78*** (4.71)
ln(mv)			-2.95*** (-9.01)	-3.24*** (-4.54)	-3.50*** (-4.69)			-3.16*** (-9.21)	-3.25*** (-3.83)	-3.15*** (-3.89)
market beta			0.25 (0.38)	0.00 (0.01)	-0.09 (-0.14)			0.35 (0.61)	-0.01 (-0.01)	0.07 (0.13)
SMB beta				-0.97** (-2.28)	-1.09** (-2.46)				-1.19** (-2.23)	-1.12** (-2.47)
HML beta				0.84** (2.26)	0.86** (2.37)				1.13*** (3.32)	1.09*** (3.22)
MOM beta				-1.25** (-2.18)	-1.09* (-1.85)				-1.41** (-2.22)	-1.26* (-1.95)
ln(ILLIQ)				-0.28 (-0.25)	-0.83 (-0.65)				0.18 (0.13)	0.37 (0.26)
turnover				-0.36** (-2.28)	-0.35** (-2.22)				-0.44** (-2.44)	-0.40** (-2.52)
% Style Ownership	-	YES	-	-	YES	-	YES	-	-	YES
OtherControls	-	-	-	-	YES	-	-	-	-	YES
Adj-R^2(%)	9.1	10.0	13.1	13.3	13.8	10.1	10.5	14.3	14.6	15.9
Number of observations	55.564	55,564	50,357	50,351	49,425	51,063	51,063	46,225	46,219	45,355

8. Summary and conclusions

In this paper we examined the relation between fund style-concentration in stock ownership with expected stock returns. It is the first paper to examine the effect of ownership concentration by institutional investors, who are nowadays the predominant investors in the market, with an average participation in individual stocks of about 80%. Fund style concentration introduces market segmentation and a varying degree of participation or inattention in the demand for stocks in a manner described by Merton (1987): Higher concentration (or lower participation) leads to lower equilibrium prices in the short-run and higher subsequent returns.

We measure style concentration in the ownership of a stock by the Herfindahl index of the shares owned through the different investment styles of institutional investors. This empirical proxy is intimately related to Merton's theoretical variable of participation. It exhibits wide cross-sectional variation among the different stocks and its cross-sectional correlation with other determinant of stock returns is relatively low.

We explore the econometric relation between style concentration in a stock during the current quarter and its return in the following quarter. Our results indicate that style concentration of ownership is positively and significantly correlated with the following quarter's stock returns. The effect is economically significant, since a one standard deviation change in style concentration predicts on average an annual return premium much higher than 2.0%.

The econometric results are robust to the presence of a host of control variables, including known stock return determinants, such as traditional risk factors or other liquidity and volatility variables. They are also robust to the presence of variables related to the various investment styles themselves, such as the percentage ownership of the stock by each specific style. And they are robust to the exclusion of the quarters of the financial crisis of 2007-2009, or the presence of outliers.

The effect of style concentration on future stock returns is present over multi-year horizons extending to four-years. The multi-year effect is both economically and statistically significant. This persistence clearly differentiates the style concentration effect, which originates from Merton's (1987) lack of participation hypothesis, from style investing effects, which are transient in their nature and originate from behavior, which is modeled in Barberis and Shleifer (2003).

Appendix A: Description of the investments styles used in the analysis

In this section we present, in alphabetical order, the different investment styles, as reported by Thomson Financial:

1) **Aggressive Growth:** Aggressive growth investors employ an extreme version of the growth style. This can be seen by their propensity to hold the stocks of companies that are growing their revenue and EPS extremely quickly, are in an early stage of their life cycle, or have minimal or no current earnings.

2) **Arbitrage:** There is not exact description in the ownership glossary of Thomson One. In this category are included all the arbitrage oriented hedge funds which are not explicitly reported as any of the following arbitrage categories: Convertible Arbitrage, Fixed Income Arbitrage, Capital Structure Arbitrage or Statistical Arbitrage.

3) **Broker Dealer:** Broker-Dealers are usually trading facilitators rather than investors. Included in this group are sell-side research firms with broker operations, NYSE and NASDAQ trading desk positions of investment banks, investment banking client desks that execute buyback programs on behalf of corporations, private client firms that essentially act as custodians for high net worth individuals, and brokers that sell unit investment trusts or exchange traded products.

4) **Capital Structure Arbitrage:** This strategy exploits the pricing inefficiencies that exist in the capital structure of the same issuer. An example is going long on a high yield bond and shorting the stock of an issuer, to hedge the equity risk component of the high yield bond.

5) **Convertible Arbitrage:** Hedge fund managers in this category construct long portfolios of corporate convertible securities, such as convertible bonds, convertible preferred stock, and warrants, and hedge the equity element of these positions by selling short some portion of the common stock into which the convertible securities may be converted.

6) **Core Growth:** Core Growth managers typically invest in mid or large capitalization, blue chip companies that have historically performed near the top of their sector or the S&P 500 in terms of profitability, earnings growth, and revenue growth. These investors are often willing to pay premium P/E multiples for highly sustainable businesses, strong management and consistent growth over the long term.

7) **Core Value:** Core Value investors focus on buying companies at relatively low valuations on an absolute basis, in relation to the market or its peers, or in comparison to an individual stock's historical levels. These portfolios typically exhibit price-to-earnings, price-to-book and price-to-cash flow multiples below the S&P 500. In addition, secular revenue growth rates of the companies in these portfolios are frequently below market averages and their earnings tend to be more cyclical.

8) **CTA/Managed Futures:** Generally trade commodity futures, financial futures, options and foreign exchange and most are sometimes highly leveraged. Traditional CTAs or trend followers attempt to capture a term trend across a range of markets.

9) **Deep Value:** Deep Value investors employ a more extreme version of value investing that is characterized by holding the stocks of companies with extremely low valuation measures. Often these companies are particularly out-of-favor or in industries that are out-of-favor. Some investors in this category are known for agitating for changes such as new management, a merger, or the spin-off of a subsidiary.

10) **Distressed Securities:** Buying and occasionally shorting securities of companies where the security's price has been, or is expected to be, affected by a distressed situation. This may involve reorganizations, bankruptcies, distressed sales and other corporate restructurings.

11) **Emerging Markets:** These investors focus primarily on companies in the developing economies of Latin America, the Far East, Europe, and Africa.

12) **Emerging Markets Hedge:** Emerging market hedge funds focus on equity or fixed income investing in emerging markets as opposed to developed markets. Emerging markets investors generally have a strong long bias.

13) **Equity Hedge:** There is not exact description in the ownership glossary of Thomson One. In this category are included all the equity oriented hedge funds which are not explicitly reported as any of the following equity hedge categories: Long / Short, Long Bias, Short Bias or Market Neutral.

14) **Event Driven:** There is not exact description in the ownership glossary of Thomson One. In this category are included all the event-driven oriented hedge funds which are not explicitly reported as any of the following event-driven categories: Merger / Risk Arbitrage or Distressed Securities.

15) **Fixed Income Arbitrage:** This trading style describes a wide variety of strategies involving fixed income securities. Hedge fund managers attempt to exploit relative mis-pricing between related sets of fixed income securities. The generic types of fixed income hedging trades include: yield curve arbitrage, corporate versus Treasury Swap yield spreads and cash versus futures.

16) **Fund of Funds:** A hedge fund which invests in other hedge funds. Funds of funds can invest in multiple managers of a single strategy or multiple strategies.

17) **GARP (Growth at a Reasonable Price):** These securities trade at a discount to the market but are expected to grow at higher than the market average. To be classified a GARP stock a company will have the following fundamentals: Forward P/E less than S&P 500 Average; and 5 Year Estimated EPS Growth greater than S&P 500 Average.

18) **Macro:** This strategy employs an opportunistic approach attempting to capitalize on global macro-economic trends across markets and sectors. This approach is primarily based on economic analysis and forecasts of shifts in interest rates, currencies, equities and commodities, as well as monetary and other public policy developments.

19) **Growth:** Growth investors bridge the gap between the Aggressive Growth and Core Growth investment styles. They tend to be slightly more aggressive than Core Growth investors, willing to pay slightly higher multiples for stocks and trade at a slightly more active pace. In general, they are looking for companies growing at superior rates than the general marketplace, but are unwilling to pay the extremely high multiples associated with the hyper growth stocks.

20) **Hedge Fund:** Hedge Fund investors have the majority of their funds invested in some sort of market neutral strategy. Notably, the term 'hedge fund' is both a legal structure (as opposed to a mutual fund) and an investment style. Nearly every firm that uses a hedge fund or market neutral style is legally organized as a hedge fund (and thus only opens to accredited investors). Many are offshore funds that are unregistered, have no investment limitations, and are not subject to disclosure regulations. The common element is that any long position taken in a specific equity is offset by a short position in either a merger partner (risk arbitrage), an 'overvalued' member of the same sector (long/short paired trading), a convertible bond (convertible arbitrage), a futures contract (index arbitrage) or an option contract (volatility arbitrage). Because of the idiosyncratic nature of these investors, the fundamentals of their portfolios are not indicative of their investment styles. Thomson Financial categorizes these portfolios based on its specific knowledge of their historical investment behavior.

21) **Income Value:** Income Value investors are similar to those in the Core Value category except they are as interested in the dividend yield as they are in the low valuation ratios of the stocks they purchase. As a result, Income Value portfolios typically exhibit above average current income and low P/E ratios.

22) **Index:** Index investors generally create portfolios that are designed to match the composition of one or more of the broad-based indices such as the S&P 500, the Russell 1000/2000/3000, the Wilshire 5000, or the NASDAQ 100. Therefore, the performance and risk of the portfolio mirrors a section of the broader market. Their investment decisions are driven solely by the makeup of the index that is tracked rather than by an

evaluation of the company and its business prospects. As a result, Index firms are often referred to as Passive investors. Thomson Financial categorizes these portfolios based on its specific knowledge of their historical investment behavior.

23) **Long / Short:** This strategy seeks to achieve absolute capital appreciation by investing in equity securities. The risk associated with long investment positions is reduced by taking short positions in securities that are thought to be overvalued.

24) **Market Neutral:** Invests in long and short equity positions. Neutrality can be established in terms of dollar exposure, beta exposure, exposure to sectors, industries, market capitalization, interest rate sensitivity, and other risk factors.

25) **Mixed Strategy:** There is not exact description in the ownership glossary of Thomson One.

26) **Momentum:** Momentum institutions invest in stocks whose price, earnings, or earnings estimates are advancing at a faster rate than the market or other stocks in the same sector. Momentum investors generally look for stocks experiencing upward earnings revisions or producing positive earnings surprises. Most of the investors in this category have relatively high portfolio turnover rates due to a short-term (often quarterly) focus, and therefore will liquidate positions at the slightest hint of a disappointment or deceleration in earnings. Thomson Financial categorizes these portfolios based on its knowledge of their historical investment behavior.

27) **Multi-Strategy:** Investment approach is diversified by employing various strategies simultaneously to realize short- and long-term gains.

28) **Quantitative / Statistical Arbitrage:** This strategy profit from pricing inefficiencies identified through the use of mathematical models.

29) **Sector Specific:** Sector Specific investors have the majority of their assets in a single major industry category. Many times these investors are "forced" to own most if not all of

the stocks in a given sector whether or not they are deemed appropriately valued. Since their portfolio exposure is linked to a single sector, their performance is usually measured against an index that is pertinent only to that industry. As such, tweaking the relative exposure to the companies that constitute a given sector will determine these firm's investment decisions.

30) **Specialty:** This category encompasses a range of styles that are not based on the fundamentals of the stocks in the portfolio relative to the overall market. Examples include investors that hold a particularly high concentration of a single stock or a very small set of stocks, or specialize in convertible securities. This category is also reserved for any institution or mutual fund that does not meet the criteria for any of the other investment styles. Thomson Financial categorizes these portfolios based on its specific knowledge of their historical investment behavior.

31) **VC/Private Equity:** Venture Capital and Private Equity investors are usually owners of public companies only when they have participated in a round of financing prior to an IPO and subsequently retained ownership after the transition from a private company to a public company. Other investors often consider positions held by venture capitalists as an "overhang" on the stock of a publicly traded company since VCs will typically dispose of their holdings of public companies during the first few years following an IPO.

32) **Yield:** Yield investors typically focus on buying companies with indicated dividend yields that are comfortably above the S&P 500 average and that are perceived to be able to continue making or increasing dividend payments over time. Investors that fall into this category tend to focus on income and safety more than on capital appreciation, and many have a dividend yield "hurdle rate" below which they will be either unlikely to consider owning a particular stock or forced to pare back a current position.

Appendix B: Mathematical relation between style concentration H and Merton's participation q

In Section 3 we showed that under the assumption that the ownership of a stock is equally divided among its owners, Merton's (1987) variable of participation (q in his paper) is equal to the inverse of our variable of ownership concentration, *H*. In this appendix, we generalize the result by relaxing the assumption of equality of the different investment shares x_j of a stock. Let j denote investor j in a particular stock, K the total number of investors in the stock, each holding a share x_j of the stock. The Herfindahl index *H* of the ownership of a stock is:

$$H = \sum_{j=1}^{K} \left(\frac{x_j}{MV}\right)^2 = \sum_{j=1}^{K} \left(\frac{\bar{x}}{MV} + \frac{x_j - \bar{x}}{MV}\right)^2 =$$

$$= \sum_{j=1}^{K} \left(\left(\frac{\bar{x}}{MV}\right)^2 + 2 \cdot \left(\frac{\bar{x}}{MV}\right) \cdot \left(\frac{x_j - \bar{x}}{MV}\right) + \left(\frac{x_j - \bar{x}}{MV}\right)^2\right) =$$

$$= \sum_{j=1}^{K} \left(\frac{\bar{x}}{MV}\right)^2 + 2 \cdot \sum_{j=1}^{K} \left(\frac{\bar{x}}{MV}\right) \cdot \left(\frac{x_j - \bar{x}}{MV}\right) + \sum_{j=1}^{K} \left(\frac{x_j - \bar{x}}{MV}\right)^2 =$$

$$= N_K \cdot \frac{\bar{x}^2}{MV^2} + \sum_{j=1}^{K} \left(\frac{x_j - \bar{x}}{MV}\right)^2 \quad (A)$$

since $\sum_{j=1}^{K}(x_j - \bar{x}) = 0$.

The total capitalization of a stock MV could also be written as $N_K \cdot \bar{x}$, which is the mean share value times the number of different investors that are present to the stock K. Then (A) becomes:

$$(A) = N_K \cdot \frac{\bar{x}^2}{{N_K}^2 \cdot \bar{x}^2} + \frac{1}{MV^2} \cdot \sum_{j=1}^{K} (x_j - \bar{x})^2 = \frac{1}{N_K} + \frac{1}{MV^2} \cdot \sum_{j=1}^{K} (x_j - \bar{x})^2,$$

which is equivalent with the Herfindahl index of the simple case of equal divided shares of the stock (which in turn coincides with the inverse of Merton's participation variable) plus a positive value which is the "variance" of the values of the shares that the shareholders hold. In the case that all the shareholders keep equal amount of shares, the $\frac{1}{MV^2} \cdot \sum_{j=1}^{K}(x_j - \bar{x})^2 = 0$ and then the $StyleConc$ is simplified to that of the simple case. On the other hand, the higher are the inequalities in the ownership shares, the higher is the "penalty" to the concentration variable.

Appendix C: Checking the robustness of the econometric results

We now run a number of additional regressions to check the robustness of our results. We begin with the question of how influential the international financial crisis was in the derivation of our econometric results. Since the international financial crisis was a very special period within the post world war II time period, we want to ensure that our results are not driven by a relatively short and abnormal time period.

Table 9 presents the earlier set of regressions in Table 6, which are now run in a smaller sample, one that excludes the 8 volatile quarters 2007-Q2 through 2009-Q1 of the international financial crisis. It turns out the regression coefficient β of style concentration *H* either stays the same or becomes stronger than before. In column 10, which includes all the control variables simultaneously, the point estimate of β becomes 23.21, which is much higher than 17.46, the corresponding estimate in Table 6. In addition, the t-statistics are also higher because of both the higher point estimates and the lower standard errors. The conclusion is that the relationship between style concentration and future stock returns is not driven by the events of the international crisis. Quite the opposite, the high volatility of that period tends to create noise, hiding rather than revealing the effect.

Next we turn to the concern we expressed earlier about the presence of outliers in the measurement of our independent variable *H*. Recall that in quite a few stocks there were times that the stock lacked participation to an extreme degree. This resulted in an extremely skewed distribution of the concentration parameter *H*, which even took values of 0.50 or higher (see Figure 3). We thus want to know whether the estimated relation between *H* and expected stock reurns is unduly influenced by the outliers in *H*.

Table 10 presents the results after winsorizing the distribution of *H* at 0.5. Namely, values of *H* larger than 0.5 are replaced with 0.5 itself, and then the regressions in Table 6 are rerun. The winsorization does not seem to change the results, except the values of the coefficients are now higher. This may be a rather expected result, which is due to the truncation of high values to the lower 0.5. The t-statistics are similar in all cases, confirming that the results of Table 6 are not driven by *H* outliers.

Next, we extend the winsorization to all the variables. We winsorize all the independent variables except the 32 investment styles in columns 2 and 10, including *H*, at the 0.5% level on both tails of their distribution. We also winsorize the dependent variable at

the 0.5% level on both tails of its distribution. This is done separately for the returns of each forecasting horizon.

Table 11 repeats the univariate and the full specified regressions (columns (1) and (10)) of Tables 6, 7 and 8. The results on the β coefficient of style concentration *H* are slightly smaller in most of the cases, compared to the basic econometric results of Tables 6, 7, and 8 (except from the univariate regression of the one-quarter horizon in which the magnitude of the coefficient is significantly smaller compared to the basic case). Yet the t-statistics tend to be substantially higher compared to those of the earlier tables. Hence we conclude that our results are not driven by the presence of outliers in any of the variables.

Table 9: Is the international financial crisis driving the results?
Sample excludes the 3Q/2007-2Q/2009 period

Panel OLS regressions of the annualized quarterly stock i return at quarter q+1, $r_{i,q+1}$, on the style concentration in ownership of stock i, $H_{i,q}$, of the previous quarter q, and on other lagged control variables for stock i, $Z_{i,q}$, which are also observed during quarter q:

$$r_{i,q+1} = \alpha + \beta \cdot (H)_{i,q} + \Gamma' \cdot Z_{i,q} + e_{i,q+1}.$$

There are 10 regressions in columns 1 through 10. A time effect with quarterly dummies is included in every regression. The variables of each regression are described in the very left column. The variables denoted as "% Style Ownership" are the ownership percentage shares of 32 different investment styles (we include only 31 of the 32 styles to avoid perfect multicollinearity). The variables denoted as "Other Controls" are the following: illiquidity beta, price momentum and the ln(dta)). See Table 1 for the detailed definitions of the variables.

Returns are measured in percentage form. The sample covers the period between 1997-Q1 and 2015-Q4, excluding the 8 quarters of the financial crisis: 2007-3Q to 2009-2Q (68 quarters) and consists of 830 stocks on average in each quarter. The total number of observations in each regression is described in the last row. t-statistics are inside the parentheses below the regression coefficients, which are based on White (1980) heteroskedasticity-consistent standard errors. Three asterisks *** denote statistical significance at the 1% level, two asterisks ** at the 5% level, and a single asterisk * at the 10% level. Adj-R^2 is the adjusted coefficient of determination of the regression, expressed in %.

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
H	**21.28*** (5.00)**	**20.18*** (4.37)**	**22.13*** (4.79)**	**21.82*** (4.74)**	**11.65** (2.52)**	**12.21*** (2.75)**	**11.88*** (2.68)**	**15.09*** (2.84)**	**17.75*** (2.86)**	**23.21*** (3.36)**
ln(mtb)						-1.54** (-2.52)	-1.47** (-2.30)	-1.66*** (-2.71)	-1.76*** (-3.34)	-2.11*** (-3.76)
ln(idio_vol)					2.66** (2.25)	2.59** (2.34)	2.68*** (2.40)	2.90*** (2.62)	2.18** (2.09)	0.80 (0.73)
ln(size)					-2.34*** (-10.24)	-2.00*** (-9.33)	-2.11*** (-9.23)	-2.69*** (-5.46)	-2.75*** (-4.93)	-2.35*** (-4.10)
market beta			0.09 (0.15)	0.19 (0.31)	-0.51 (-0.86)	-0.23 (-0.41)	-0.29 (-0.50)	-0.22 (-0.40)	-0.63 (-1.13)	-0.90 (-1.62)
SMB beta				0.71* (1.87)			-0.54 (-1.39)		-0.61 (-1.60)	-0.70* (-1.83)
HML beta				0.55 (1.39)			-0.06 (-0.16)		0.09 (0.26)	0.18 (0.51)
MOM beta				-0.83 (-1.52)			-0.66 (-1.31)		-0.69 (-1.36)	-0.66 (-1.29)
ln(ILLIQ)								-1.48 (-1.64)	-1.56 (-1.49)	-1.41 (-1.27)
turnover								-0.70*** (-4.25)	-0.70*** (-4.30)	-0.72*** (-4.23)
% of Style Ownership	-	YES	-	-	-	-	-	-	-	YES
Other Controls	-	-	-	-	-	-	-	-	YES	YES
Adj-R^2 (%)	18.3	18.6	18.2	18.2	18.6	19.6	19.6	19.7	20.4	20.5
Number of observations	62,481	62,481	60,273	60,273	59,365	57,436	57,436	57,431	56,424	56,424

Table 10: How important are the outliers? Style concentration *H* winsorized at 0.50

Panel OLS regressions of the annualized quarterly stock i return at quarter q+1, $r_{i,q+1}$, on the style concentration in ownership of stock i, $H_{i,q}$, of the previous quarter q, which is winsorized at 0.5 for values higher than 0.5 (i.e., for those values, the number 0.5 is used), and on other lagged control variables for stock i, $Z_{i,q}$, which are also observed during quarter q:

$$r_{i,q+1} = \alpha + \beta \cdot (H)_{i,q} + \Gamma' \cdot Z_{i,q} + e_{i,q+1}.$$

There are 10 regressions in columns 1 through 10. A time effect with quarterly dummies is included in every regression. The variables of each regression are described in the very left column. The variables denoted as "% Style Ownership" are the ownership percentage shares of 32 different investment styles (we include only 31 of the 32 styles to avoid perfect multicollinearity). The variables denoted as "Other Controls" are the following: illiquidity beta, price momentum and the ln(dta)). See Table 1 for the detailed definitions of the variables.

Returns are measured in percentage form. The sample covers the period between 1997-Q1 and 2015-Q4, (76 quarters) and consists of 838 stocks on average in each quarter. The total number of observations in each regression is described in the last row and is identical to those in Table 6. t-statistics are inside the parentheses below the regression coefficients, which are based on White (1980) heteroskedasticity-consistent standard errors. Three asterisks *** denote statistical significance at the 1% level, two asterisks ** at the 5% level, and a single asterisk * at the 10% level. Adj-R^2 is the adjusted coefficient of determination of the regression, expressed in %.

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
H	**35.51*** (6.07)**	**31.11*** (4.95)**	**38.10*** (5.98)**	**34.60*** (5.21)**	**15.93** (2.54)**	**16.27*** (2.68)**	**15.41** (2.55)**	**20.87*** (2.82)**	**19.97*** (2.59)**	**19.99** (2.34)**
ln(mtb)						-3.57*** (-5.32)	-3.32*** (-4.83)	-3.69*** (-5.50)	-3.16*** (-5.10)	-3.64*** (-5.47)
ln(idio_vol)					6.10*** (5.13)	6.10*** (5.36)	6.30*** (5.38)	6.47*** (5.67)	6.69*** (5.77)	5.34*** (4.42)
ln(mv)					-2.62*** (-11.69)	-1.92*** (-8.75)	-2.12*** (-8.91)	-2.71*** (-5.47)	-2.99*** (-5.28)	-2.58*** (-4.45)
market beta			1.15* (1.94)	0.61 (0.99)	0.20 (0.34)	0.27 (0.46)	0.07 (0.12)	0.28 (0.48)	-0.14 (-0.25)	-0.37 (-0.64)
SMB beta				0.46 (1.10)			-1.02** (-2.20)		-1.09** (-2.35)	-1.15** (-2.47)
HML beta				1.12*** (2.94)			0.25 (0.66)		0.38 (1.09)	0.44 (1.28)
MOM beta				-1.71*** (-3.16)			-1.09** (-2.17)		-1.06** (-2.11)	-1.04** (-2.05)
ln(ILLIQ)								-1.68* (-1.86)	-1.89* (-1.79)	-1.34 (-1.20)
turnover								-0.76*** (-4.41)	-0.75*** (-4.39)	-0.78*** (-4.36)
% of style Ownership	-	YES	-	-	-	-	-	-	-	YES
Other Controls	-	-	-	-	-	-	-	-	YES	YES
Adj-R^2 (%)	21.3	21.5	21.6	21.3	21.8	23.0	23.0	23.0	23. 0	23. 7
Number of observations	70,490	70,490	67,881	67,881	66,971	64,807	64,807	64,802	63,704	63,704

Table 11: How important are the outliers? Winsorizing all variables at the 0.5% level at each tail of their distribution

Panel OLS regressions of the annualized stock i return $r_{i,q+k}$ from the end of quarter q to the end of quarter q+k, where k = either 1 (columns (1)-(2)), 4 (columns (3)-(4)), 8 (columns (5)-(6)), 12 (columns (7)-(8)) or 16 (columns (9)-(10)), on the style concentration in ownership of stock i, $H_{i,q}$, of quarter q, and on other lagged control variables for stock i, $Z_{i,q}$, which are also observed during quarter q:

$$r_{i,q+1} = \alpha + \beta \cdot (H)_{i,q} + \Gamma' \cdot Z_{i,q} + e_{i,q+1}.$$

All the dependent and independent variables are winsorized at the 0.5% of each tail, except for the 32 variables: % of style ownership. There are 10 regressions in columns 1 through 10. A time effect with quarterly dummies is included in every regression. The variables of each regression are described in the very left column. The variables denoted as "% Style Ownership" are the ownership percentage shares of 32 different investment styles (we include only 31 of the 32 styles to avoid perfect multicollinearity). The variables denoted as "Other Controls" are the following: illiquidity beta, price momentum and the ln(dta)). See Table 1 for the detailed definitions of the variables. Returns are measured in percentage form. The sample covers the period from of 1997-Q1 to 2016-Q1 (for k=1), or 2015-Q1 (for k = 4) or 2014-Q1 (for k = 8) or 2013-Q1 (for k=12) or 2012-Q1 (for k=16). The quarterly cross section consists on average of around 838 stocks in the one-quarter horizon, 818 stocks in the one-year horizon, 794 stocks in the two-year horizon, 772 in the three-year horizon, and 756 stocks in the four-year horizon. The total number of observations in each regression is described in the last row and is identical to those in Table 6. t-statistics are inside the parentheses below the regression coefficients, which are based on White (1980) heteroskedasticity-consistent standard errors for the one-quarter horizon and Newey-West (1987) for the longer horizons. Three asterisks *** denote statistical significance at the 1% level, two asterisks ** at the 5% level, and a single asterisk * at the 10% level. Adj-R^2 is the adjusted coefficient of determination of the regression, expressed in %.

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
	1q-Gross Returns	**1q- Gross Returns**	**1y- Gross Returns**	**1y- Gross Returns**	**2y- Gross Returns**	**2y- Gross Returns**	**3y- Gross Returns**	**3y- Gross Returns**	**4y- Gross Returns**	**4y- Gross Returns**
H	**14.58*** (4.37)**	**14.09** (2.49)**	**20.42*** (6.64)**	**18.76*** (3.62)**	**23.13*** (6.61)**	**20.12*** (3.58)**	**23.39*** (6.12)**	**17.47*** (2.81)**	**26.67*** (5.68)**	**15.59** (2.12)**
ln(mtb)	-	-3.00*** (-5.68)	-	-1.35*** (-2.85)	-	-1.33*** (-2.66)	-	-1.48*** (-2.68)	-	-2.33*** (-3.68)
ln(idio_vol)	-	4.51*** (4.31)	-	4.87*** (5.48)	-	4.71*** (5.35)	-	4.58*** (5.19)	-	6.49*** (6.88)
ln(mv)	-	-3.73*** (-6.52)	-	-3.83*** (-7.54)	-	-4.09*** (-7.75)	-	-4.09*** (-6.82)	-	-3.96*** (-5.84)
market beta	-	0.21 (0.45)	-	1.34*** (3.21)	-	1.04** (2.47)	-	0.34 (0.81)	-	0.26 (0.59)
SMB beta	-	-0.43 (-1.22)	-	-0.80*** (-2.67)	-	-0.39 (-1.29)	-	-0.54* (-1.70)	-	-0.56 (-1.61)
HML beta	-	0.57* (1.85)	-	1.13*** (4.08)	-	1.01*** (3.56)	-	1.15*** (4.19)	-	1.26*** (4.45)
MOM beta	-	-0.61 (-1.40)	-	-0.42 (-1.10)	-	-0.11 (-0.26)	-	-0.74* (-1.81)	-	-0.87** (-2.04)
ln(ILLIQ)	-	-5.18*** (-4.41)	-	-4.40*** (-4.31)	-	-4.42*** (-4.10)	-	-3.70*** (-3.19)	-	-3.11** (-2.39)
turnover	-	-5.00*** (-5.04)	-	-4.56*** (-5.69)	-	-3.52*** (-4.51)	-	-3.93*** (-5.05)	-	-4.36*** (-5.23)
% of style Ownership	-	YES	-	YES	-	YES	-	YES	-	YES
Other Controls	-	YES	-	YES	-	YES	-	YES	-	YES
Adj-R^2(%)	24.2	25.6	20.4	23.2	18.9	23.0	15.0	20.2	14.3	21.1
Number of observations	70,490	63,704	65,589	58,889	60,466	53,960	55,564	49,425	51,063	63,704

Appendix D: Style concentration vs. style investing

We further check the robustness of our results against the effects of style investing, with the addition of the lagged style returns as independent variables to our econometric analysis. Following the "style box" of Morningstar, for each month from 1/1995 to 12/2015, we distribute the stocks of the sample to the following 9 styles:

1) Small – Value (the size of the stock below the 30th percentile of the NYSE stocks and its book-to-market (the inverse of the market-to-book variable) above the 70th percentile of the NYSE stocks).[36]
2) Small – Blend (the size of the stock below the 30th percentile of the NYSE stocks and its book-to-market between the 30th and the 70th percentile of the NYSE stocks).
3) Small – Growth (the size of the stock below the 30th percentile of the NYSE stocks and its book-to-market below the 30th percentile of the NYSE stocks).
4) Mid-Cap – Value (the size of the stock between the 30th and the 70th percentile of the NYSE stocks and its book-to-market above the 70th percentile of the NYSE stocks).
5) Mid-Cap – Blend (the size of the stock between the 30th and the 70th percentile of the NYSE stocks and its book-to-market between the 30th and the 70th percentile of the NYSE stocks).
6) Mid-Cap – Growth (the size of the stock between the 30th and the 70th percentile of the NYSE stocks and its book-to-market below the 30th percentile of the NYSE stocks).
7) Big – Value (the size of the stock above the 70th percentile of the NYSE stocks and its book-to-market above the 70th percentile of the NYSE stocks).
8) Big – Blend (the size of the stock above the 70th percentile of the NYSE stocks and its book-to-market between the 30th and the 70th percentile of the NYSE stocks).
9) Big – Growth (the size of the stock above the 70th percentile of the NYSE stocks and its book-to-market below the 30th percentile of the NYSE stocks).

The above classification is done for each stock separately every month. It does not coincide exactly with the styles reported by Thomson Financial, which we used earlier for the measurement of style concentration. However, the criteria that are used by Thomson Financial are similar with the criteria that we use to create the 9 different styles. After all,

[36] We use the breakpoints that are provided at the electronic library of Kenneth French.

size and book-to-market are the most used variables in the determination of the majority of styles. Hence our methodology promises to capture a large part of the style investing effects.

Next, we estimate the monthly style return for each of the 9 different styles, as the equally-weighted average of the monthly returns of the stocks belonging to the corresponding style, at the specific month. We thus create 9 time-series of style returns from 1/1995 to 12/2015.[37] We subsequently calculate the quarterly style returns, using the appropriate compounding. The nine quarterly time series of the styles will be subsequently used to draw data for the regressions.

Since our sample frequency is quarterly, we need to classify a stock as belonging into a particular investment style every quarter. We use the classification of the last (third) month of the quarter q to characterize the full quarter. Once we have determined the style of the stock for quarter q, we use its style's lagged quarterly returns as additional control variables in the regressions.

The above approach is similar to that of Teo and Woo (2004) and Froot and Teo (2008), who examine the effect of past style returns to the future stock returns. Their papers confirm empirically the style investing theory of Barberis and Shleifer (2003). Teo and Woo (2004) find that style returns of the past quarter positively predict future monthly stock returns, while style returns of the past year negatively predict future monthly stock returns. This is explained as a reversal of prices towards equilibrium, after an initial shock due to style investing. In addition, Froot and Teo (2008) show that at weekly frequencies, style returns positively predict a transitory component of future stock returns. They also show that this effect weakens over time and fully dissipates after several weeks.

Table 12 presents the results. It includes five forecasting horizons: 1-quarter, 1-year, 2-years, 3-years and 4-years. In each horizon, there are two regressions, which are extensions of the univariate case and the full specification case with all previous control variables of the earlier tables. The extra variables now, are four lags of the quarterly style returns, as described above.

The results in Table 12 are in line with the empirical findings of the style investing literature. Past style returns of the immediate previous quarter positively predict the stock

[37] The breakpoints of BE/ME are annual and are available until 2015, thus we could classify the stocks and create the style returns only until the last quarter of 2015. This fact does not affect our analysis, since the last observations of the independent variables are measured at the 4Q2015.

returns of the following quarter (columns (1) and (2) of Table 12). However this is not the case for the more distant quarterly lags, since the style returns lagged 2 or 4 quarters predict negatively the future stock returns. This is not a surprising result, as Teo and Woo (2004) also find that the positive effect of past style returns takes place in short horizons, while Froot and Teo (2008) find strong positive relation between past style returns and future stock returns, on weekly frequency. In addition, this is evidence of reversal of the style effect on future stock returns, in line with the results of Teo and Woo (2004).

The predictability of the lagged past style returns change sign (from positive to negative) in most of the cases in the longer horizons of one to four years (columns (3)-(10) of Table 12). The magnitudes of the coefficients of the lagged style returns become much smaller at those longer horizons, indicating that the effects gradually dissipate. Overall, these findings underpin the theoretical predictions of Barberis and Shleifer (2003), that the prices of stocks that belong to styles with positive past returns, increase, and subsequently decrease in longer horizons, towards their equilibrium level.

In all the regressions of Table 12, the coefficient β of style concentration *H* remains positive and significant (except for the case of 4-years ahead, where the t-statistic equals 1.59, a bit lower than 1.74 in Table 7). These findings further confirm the earlier conclusion that the effect of style concentration is an equilibrium effect, which is distinct from the transient effects of style investing.

Table 12: Inclusion of quarterly lagged style returns as further controls to the regressions in Tables 6, 7, 8

Panel OLS regressions of the annualized stock i return $r_{i,q+k}$ from the end of quarter q to the end of quarter q+k, where k = either 1, 4, 8, 12, 16, on the style concentration in ownership of stock i, $H_{i,q}$, of quarter q, and on other control variables for stock i, $Z_{i,q}$, which are also observed during quarter q:

$$r_{i,q+k} = \alpha + \beta \cdot (H)_{i,q} + \Gamma' \cdot Z_{i,q} + e_{i,q+k}.$$

There are 10 regressions in columns 1 through 10. A time effect with quarterly dummies is included in every regression. The variables of each regression are described in the very left column. The variables denoted as "% Style Ownership" are the ownership percentage shares of 32 different investment styles (we include only 31 of the 32 styles to avoid perfect multicollinearity). The variables denoted as "Other Controls" are the following: market beta, SMB beta, HML beta, MOM beta, ln(ILLIQ), turnover, illiquidity beta, price momentum and the ln(dta)). See Table 1 for the detailed definitions of the variables.

The new variables in Table 12 are Style ret 1q lagged, ..., Style ret 4q lagged. Each stock in quarter q belongs to a particular style. We assign to the stock in quarter q, the lags 1 to 4 of its own style.

Returns are measured in percentage form. The sample covers the period between 1997-Q1 and 2015-Q4 (76 quarters) and on average consists of around 838 stocks in each quarter. The total number of observations in each regression is described in the last row. t-statistics are inside the parentheses below the regression coefficients, which are based on White (1980) heteroskedasticity-consistent standard errors for the columns (1) and (2) and on Newey-West (1987) for the columns (3) to (10). Three asterisks *** denote statistical significance at the 1% level, two asterisks ** at the 5% level, and a single asterisk * at the 10% level. Adj-R^2 is the adjusted coefficient of determination of the regression, expressed in %.

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
	1q-Raw Returns	**1q-Raw Returns**	**1y-Raw Returns**	**1y-Raw Returns**	**2y-Raw Returns**	**2y-Raw Returns**	**3y-Raw Returns**	**3y-Raw Returns**	**4y-Raw Returns**	**4y-Raw Returns**
H	**22.00*** (5.01)**	**17.20** (2.41)**	**27.14*** (6.08)**	**17.28*** (2.66)**	**28.25*** (5.36)**	**16.33** (2.13)**	**28.77*** (5.01)**	**15.67* (1.80)**	**35.46** (4.89)**	**15.60 (1.59)**
ln(mtb)	-	-3.09*** (-3.90)	-	-0.55 (-0.91)	-	-0.90 (-1.29)	-	-1.51** (-2.28)	-	-2.29*** (-3.17)
ln(idio_vol)	-	5.38*** (4.42)	-	4.96*** (4.60)	-	5.80*** (4.37)	-	4.78*** (3.28)	-	5.92*** (3.90)
ln(mv)	-	-2.53*** (-4.33)	-	-2.62*** (-4.93)	-	-3.39*** (-5.42)	-	-3.31*** (-4.55)	-	-3.09*** (-3.84)
style ret 1q lagged	25.25** (2.47)	36.95*** (3.33)	-43.87*** (-6.86)	-35.07*** (-5.90)	-40.25*** (-8.81)	-33.38*** (-7.45)	-14.66*** (-3.33)	-5.45 (-1.23)	-17.33*** (-4.74)	-9.08** (-2.55)
style ret 2q lagged	-91.41*** (-9.57)	-79.84*** (-8.46)	5.63 (1.17)	12.76*** (2.58)	-2.85 (-0.53)	7.30 (1.19)	-16.03*** (-4.71)	-4.23 (-0.90)	-9.54*** (-2.72)	7.70 (1.63)
style ret 3q lagged	30.41*** (3.29)	36.89*** (3.96)	-30.15*** (-5.27)	-23.36*** (-4.31)	-18.22*** (-3.71)	-7.39 (-1.28)	-13.61*** (-3.72)	-3.39 (-0.68)	-13.88*** (-3.40)	3.16 (0.67)
style ret 4q lagged	-18.01** (-2.09)	-12.69 (-1.43)	-4.89 (-0.95)	0.45 (0.09)	4.67 (0.36)	13.35** (2.50)	-0.65 (-0.18)	10.37** (2.14)	-16.74*** (-4.34)	-4.08 (-0.86)
% Style Ownership	-	YES	-	YES	-	YES	-	YES	-	YES
Other Controls	-	YES	-	YES	-	YES	-	YES	-	YES
Adj-R^2 (%)	22.8	23.9	17.7	20.1	13.3	15.7	10.6	13.7	10.9	15.7
Number of observations	66,012	62,924	60,957	58,112	55,897	53,193	51,216	48,753	46,923	44,685

References

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

Agarwal, P., 2010. Institutional ownership and stock liquidity. SSRN working paper.

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

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

Amihud, Y., Mendelson, H., Uno, J., 1999. Number of Shareholders and Stock Prices: Evidence from Japan. The Journal of Finance 54, 1169-1184

Ang, A., Hodrick, R.J., Xing, Y., Zhang, X., 2006. The Cross-Section of Volatility and Expected Returns. The Journal of Finance 61, 259-299

Arbel, A., Carvell, S., Strebel, P., 1983. Giraffes, Institutions and Neglected Firms. Financial Analysts Journal, 2-8

Avramov, D., Chordia, T., 2006. Asset Pricing Models and Financial Market Anomalies. Review of Financial Studies 19, 1001-1040

Badrinath, S.G., Wahal, S., 2002. Momentum Trading by Institutions. The Journal of Finance 57, 2449-2478

Ball, R., 1978. Anomalies in relationships between securities' yields and yield-surrogates. Journal of Financial Economics 6, 103-126

Banz, R.W., 1981. The relationship between return and market value of common stocks. Journal of Financial Economics 9, 3-18

Barabanov, S., McNamara, M., 2002. Market perception of information asymmetry: Concentration of ownership by different types of institutions and bid-ask spread. SSRN working paper.

Barberis, N., Shleifer, A., 2003. Style investing. Journal of Financial Economics 68, 161-199

Basu, S., 1977. Investment Performance Of Common Stocks In Relation To Their Price-Earnings Ratios: A Test Of The Efficient Market Hypothesis. The Journal of Finance 32, 663-682

Bennett, J.A., Sias, R.W., Starks, L.T., 2003. Greener Pastures and the Impact of Dynamic Institutional Preferences. Review of Financial Studies 16, 1203-1238

Bhandari, L.C., 1988. Debt/Equity Ratio and Expected Common Stock Returns: Empirical Evidence. The Journal of Finance 43, 507-528

Boyer, Brian H., 2010. Style-related Comovement: Fundamentals or Labels? Working Paper, Bringham Young University, August 4.

Brennan, M.J., Chordia, T., Subrahmanyam, A., Tong Qing 2012. Sell-Order Liquidity and the Cross-Section of Expected Stock Returns. Journal of Financial Economics 105, 523-541

Brennan, M.J., Chordia, T., Subrahmanyam, A., 1998. Alternative factor specifications, security characteristics, and the cross-section of expected stock returns. Journal of Financial Economics 49, 345-373

Brennan, M.J., Subrahmanyam, A., 1996. Market microstructure and asset pricing: On the compensation for illiquidity in stock returns. Journal of Financial Economics 41, 441-464

Brown, S.J., Goetzmann, W.N., 1997. Mutual fund styles. Journal of Financial Economics 43, 373-399

Carhart, M.M., 1997. On Persistence in Mutual Fund Performance. The Journal of Finance 52, 57-82

Chan, L.K.C., Chen, H.-L., Lakonishok, J., 2002. On Mutual Fund Investment Styles. Review of Financial Studies 15, 1407-1437

Cochrane, J.H., 2001. Asset Pricing. Princeton, NJ: Princeton University Press

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

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

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

Froot, K., Teo, M., 2008. Style Investing and Institutional Investors. Journal of Financial and Quantitative Analysis 43, 883-906

Fu, F., 2009. Idiosyncratic risk and the cross-section of expected stock returns. Journal of Financial Economics 91, 24-37

Gompers, P.A., Metrick, A., 2001. Institutional Investors and Equity Prices. The Quarterly Journal of Economics 116, 229-259

Greenwood, R., Thesmar, D., 2011. Stock price fragility. Journal of Financial Economics 102, 471-490

Grinblatt, M., Titman, S., Wermers, R., 1995. Momentum Investment Strategies, Portfolio Performance, and Herding: A Study of Mutual Fund Behavior. The American Economic Review 85, 1088-1105

Jegadeesh, N., 1990. Evidence of Predictable Behavior of Security Returns. The Journal of Finance 45, 881-898

Jegadeesh, N., Titman, S., 1993. Returns to Buying Winners and Selling Losers: Implications for Stock Market Efficiency. The Journal of Finance 48, 65-91

Lehmann, B.N., 1990. Residual risk revisited. Journal of Econometrics 45, 71-97

Lintner, J., 1965. Security Prices, Risk, and Maximal Gains From Diversification. The Journal of Finance 20, 587-615

Ľuboš Pástor, Robert F. Stambaugh, 2003. Liquidity Risk and Expected Stock Returns. Journal of Political Economy 111, 642-685

Merton, R.C., 1987. A Simple Model of Capital Market Equilibrium with Incomplete Information. The Journal of Finance 42, 483-510

Mossin, J., 1966. Equilibrium in a Capital Asset Market. Econometrica 34, 768-783

Newey, W.K., West, K.D., 1987, A Simple, Positive Semi-Definite, Heteroskedasticity and Autocorrelation Consistent Covariance Matrix. Econometrica 55, 703-708

Nofsinger, J.R., Sias, R.W., 1999. Herding and Feedback Trading by Institutional and Individual Investors. The Journal of Finance 54, 2263-2295

Petersen, M.A., 2009. Estimating Standard Errors in Finance Panel Data Sets: Comparing Approaches. Review of Financial Studies 22, 435-480

Piotroski, J.D., 2000. Value Investing: The Use of Historical Financial Statement Information to Separate Winners from Losers. Journal of Accounting Research 38, 1-41

Sharpe, W., 1964. Capital Asset Prices: A Theory of Market Equilibrium Under Conditions of Risk. The Journal of Finance 19, 425-442

Shleifer, A., Vishny, R.W., 1997. The Limits of Arbitrage. The Journal of Finance 52, 35-55

Teo, M., Woo, S.-J., 2004. Style effects in the cross-section of stock returns. Journal of Financial Economics 74, 367-398

Tinic, S.M., West, R.R., 1986. Risk, Return, and Equilibrium: A Revisit. Journal of Political Economy 94, 126-147

Wahal, S., Yavuz, M.D., 2013. Style investing, comovement and return predictability. Journal of Financial Economics 107, 136-154

Wermers, R., 1999. Mutual Fund Herding and the Impact on Stock Prices. The Journal of Finance 54, 581-622

White, H.(1980). A heteroskedasticity-consistent covariance matrix estimator and a direct test of heteroskedasticity. *Econometrica* 48(4), 817-838.