ESSAYS ON ASSET PRICING

Ph.D. Thesis

Georgios I. Karalas

Submitted in fulfillment of the requirements
for the Degree of Doctor of Philosophy

Department of Banking and Financial Management
University of Piraeus

Piraeus 2017

To my parents Ilias and Maro

and to my brother Andreas

Acknowledgements

1. I would like to thank the Department of Banking and Financial Management of the University of Piraeus for all the valuable resources provided to me for the completion of this thesis. The continuous efforts to maintain high academic standards is a significant achievement, particularly given the difficult economic environment in the last few years in our country.

2. I would like to thank my supervisor, Professor Gikas Hardouvelis for his guidance. As a true mentor, he encouraged me to bring forward real and meaningful questions as topics for my thesis and allowed me the "space" to look for them during the initial stages of my research, which I consider today a key part of my development as a researcher. As he has notably pointed out to me at the beginning of my PhD: "I will not encourage you to add to my research, but rather find a topic that you like, as this is the only way you will do it well". Now, I only fully appreciate this.

3. I would like to thank the 7-member Committee, as well as all the Faculty members of the Department for their valuable support all these years. It is very important for a young researcher to grow in an environment where research is highly regarded and appreciated.

4. I would also like to thank the Department's personnel as well as my fellow PhD students.

5. Without the support of my friends, the completion of this thesis would not have been possible. As faithful co-travellers, they supported me and kept me strong and focused on my research.

6. I would like to thank my partner Jessica for her support, patience and faith in me, which was crucial for the completion of my research.

7. Finally, I would like to thank my family and especially my parents and brother. They provided me with the qualifications and capabilities as well as the ethos and unconditional love for knowledge. Without their support and tolerance the completion of this thesis would not have been possible. I dedicate this thesis to them.

Contents

Introduction

Classical asset pricing theory assumes that the agents in the economy are identical, thus in essence there is only a representative agent. The basic asset pricing formulas - like the Capital Asset Pricing Model (CAMP) – are derived under the assumption of the existence of a representative agent. Market portfolio is mean-variance efficient and only the exposure of a stock to un-diversified systematic risk creates persistent cross-sectional differences in stock prices and in expected stock returns. However, as long as there is heterogeneity among agents in the economy and this heterogeneity creates incomplete participation in the market, deviations emerge from the classical models. In an environment in which investors hold portfolios from subsets of stocks, the classical CAPM is modified in two ways[1]: (a) the market portfolio is no more mean-variance efficient, and (b) market beta is not the only pricing factor in the cross-section of stocks. The idiosyncratic volatility, the market size of the stock and the degree of participation in the demand for a stock are additional pricing factors (Merton's Presidential Address, 1987).

The presence of non-representative agents with their limited participation in the stock market affects the demand for stocks through a second channel: the limits-to-arbitrage. One key premise of the classical asset pricing models is that demand curves for stocks (and generally for financial assets) are flat, that is, the non-fundamental demand for stocks does not affect their prices and their expected returns. Non-fundamental demand is offset by arbitrage activity and only the fundamentals of a stock determine its price and its expected return. Limited participation reduces the amount of capital available for arbitrage and, as a result, deviations in non-fundamental demand are not being fully offset and end up affecting stock prices and expected returns. Since deviations from fundamentals are temporary, the corresponding deviations in stock prices are also temporary, creating idiosyncratic stock price volatility.

---

[1] The same modification holds also for multi-factor pricing models like the 3-factor model of Fama and French. Multi-factor models add new marketwide risk-factors to capture various forms of systematic risk, thus they use more betas for the pricing of stocks. The idiosyncratic stock characteristics (size, idiosyncratic volatility, degree of participation) could be determinants of stock returns when there is limited participation and not perfect diversified portfolios. The nature of the betas of the additional risk factors is fundamentally different from that of the rest of the stock characteristics.

In real economic life, only a small fraction of individual (retail) agents participate in the stock market and even the participating ones do not hold the full portfolio of existing stocks. Limited information, limited processing abilities, limited time, limited trust, limited "rationality" and limited wealth are some of the reasons that hold back investors from forming portfolios that include all the available stocks. In real world a number of complicated schemes and investment institutions exist, which fill the gap between the economists' ideal perfect market and the real market consisting of fewer participating agents. These investment institutions (*e.g.* among others, mutual funds, index funds, hedge funds, venture capital, and others) take advantage of economies of scale and economies of scope, in order to collect superior information and to process it fast, cheaply and without behavioral biases. Institutional participation in the stock market has risen over time to represent 80% of total stock market participation today.

Investment institutions (henceforth funds), however, are not free from various forms of restrictions. First, agency problems emerge between funds and individual (retail) investors, since the latter do not know or can spend the time to learn in great detail the investment activities of the former. Individual investors need to trust the institutions and to be able to evaluate in an easy way their investment performance. Second, investment funds are themselves companies aiming to make a profit, thus they compete against each other to attract investors. In this competitive environment, information gets clouded and it becomes particularly difficult for retail investors to choose the best fund manager or the best fund, which matches their investment preferences. Third, the economies of scale existing in the formation of funds are not powerful enough or adequate to enable those funds to gather and process fully the total information for all the available stocks. Finally, funds face several constraints regarding their investment strategies (*e.g.* short sales restrictions, disclosures of their positions, geographical restrictions, leverage restrictions, *etc*) which put further restrictions to their investment behavior.

One way for institutional investors (funds) to respond to all the above restrictions and simplify the set of choices for retail investors, is the adoption of the practice of following specific investment styles or different forms of investment types. Some investment styles are formed according to simple quantitative rules that refer to some basic magnitudes of stocks (*e.g.* among others, capitalization, the ratio of the book value of a stock to its market capitalization, and combinations thereof). Stocks are then classified to these styles depending on their relative point on the cross-sectional distribution of relative characteristics. Other investment styles are based on specific investment strategies (*e.g.* index funds, which track a specific index). In addition, some funds operate under a different regulatory framework (e.g. hedge funds) avoiding legal restrictions and other restrictions that mutual funds face. The decision by an investment fund to belong to a particular investment style and often advertize it, alleviates a large number of the aforementioned fund investment strategy constraints.

Due to its simplistic approach, style separation is easily understandable for the majority of investors. This understanding mitigates the first of the previous mentioned problems of the funds (that is the agency problems that result due to asymmetric information), since it is easier for investors to monitor the strategy of the funds. In addition, the existence of certain styles makes easier the comparison between the performances of those investment institutions, hence alleviating further the agency problems stemming from moral hazard. Trust between investors and institutions may also be enhanced if the agency problems are reduced. Secondly, with the adoption of a specific investment style, an institution can advertise its services much easier; it is worth mentioning that a large number of funds add to their name the style that they follow. The easier advertisement reduces the informational costs of investors to find the institution of their choice.

An additional set of reasons for the suitability of the adoption of style separation by the funds stems from the facilitation of their own investment activities. Funds are able to limit their focus to a subset of stocks that fall within their style. In this way they reduce costs for gathering and processing information. In addition, they develop expertise to certain stocks that fall within their style. They also have a specific investment strategy, thus they limit their cost of investment planning. As a result, they are able to take advantage of their superior information and make information-intense investments, increasing the informational content of the market, and providing liquidity (reducing the severity of the third of the abovementioned funds' problems, namely the inadequate economies of scale of the funds). Moreover, special "styles" like hedge funds, are free from restrictions like short-selling and they could follow more "exotic" combinations of long-short strategies, further improving the efficiency of the market (alleviating the fourth of the previous mentioned problems about the operation of funds, the regulatory restrictions).

The adoption of investment styles does not eliminate the initial problem of limited participation in the stock market by individual (retail) investors. It simply transforms the limited participation problem of retail investors or the inattention of retail investors into a style inattention problem, this time at the fund level. This happens because the fundamental reasons for the inattention (e.g. among others, limited information, limited processing abilities, limited rationality, see paragraph 3), although reduced, are still present. The formation of funds transforms these problems from the level of the stock selection by the individual investor to two different levels; the level of the fund manager selection by the individual investor and the level of the stock selection by the fund manager. The adoption of investment styles and types facilitates both levels, as described in the two previous paragraphs.

In reality stock institutional ownership is deeply segmented according to various investment types and styles. This segmentation could affect stock prices and returns both through the classical equilibrium - premium approach and through the limits-to-arbitrage approach. Merton (1987) provides the equilibrium theoretical framework for the relation of inattention (or limited participation) to a return risk premium. The explanation is that as we move further away from the theoretical assumption of complete participation (perfect diversification) investors are exposed not only to systematic risk but to idiosyncratic risk as well. As the diversification becomes more incomplete (the participation becomes less perfect, i.e. we move further away from the CAPM world) investors become even more exposed to idiosyncratic risk and hence the level of participation demands a premium.[2]A manifestation of Merton's description about limited participation and inattention is the style-related segmentation, which – for this reason - is a possible determinant of expected stock returns. Stocks which are neglected from the majority of styles and types are further away from the complete participation world of CAPM and thus they gain a return premium.

The theoretical predictions of Barberis and Shleifer (2003), and the subsequent empirical tests of these predictions, examine the effect of style investing through the theoretical perspective of limits-to-arbitrage. Non-fundamental demand due to style investors who chase the better performing style for each period, leads to stock price changes which subsequently reverse to their fundamental levels, creating style momentum, reversals and stock return predictability. This approach is complementary to the equilibrium approach

[2] According to Merton's paper, if the participation is incomplete, the expected stock returns are a function not only of the market beta (exposure to systematic risk) but also of the triple product of the level of participation with the idiosyncratic volatility of the stock and the market size of the stock. Lower participation, higher idiosyncratic volatility and higher size predict higher expected returns.

described by Merton, since this describes transient effects, while Merton's model describes the formation of the equilibrium expected return. The two theoretical approaches are complementary and could be in work simultaneously.

The first two chapters of the dissertation take advantage of the known information about the investment positions, types and styles of institutional investors and quantify the exposure to stylization of each stock. The existing literature about style investing focuses on the stocks characterizing them according to styles. It, thus, assumes that all the stocks are fully exposed to the stylization of the market and that each stock belongs only to one style. One primary innovation-contribution of this thesis is the examination of style investing through institutional stock ownership. Instead of characterizing each stock with a single style, we characterize each fund with a single investment style, thus allowing the exposure of stocks to the continuum of fund-styles. This is closer to economic reality as individual stocks do not belong to a unique style, but they are rather held by a number of different styles.

Our new approach of measuring the exposure of each stock to style investing (through ownership), enables us to examine previous unexplored research questions regarding the effect of style investing on stock returns, stock volatility and stock liquidity. First, the measure of stylization – which is introduced in this thesis – is utilized in Chapter 1 to test for the existence of a return premium related to the stylization of stocks. That is a different perspective from that of Barberis and Shleifer (2003), since it is based on the equilibrium mechanism of Merton (predicting risk related return premium) and not in the limits-to-arbitrage theory (which predicts deviations from the fundamental price). Second, the measure of stylization is used in Chapter 2 for the examination of the – previous unexplored - relation between style investing and stock price volatility and liquidity. The relation between style investing and volatility and liquidity is based on the limits-to-arbitrage theory. Stocks which are more exposed to stylization, exhibit higher price changes from the non-fundamental demand created by style investors, and should thus exhibit higher levels of daily volatility and lower levels of liquidity.

The results of Chapter 1 confirm that lower style attention leads to higher expected stock returns. This is the first study showing that style investing does affect stock returns not only as a deviation from fundamentals, but also as an equilibrium premium. Moreover, the contribution of Chapter 1 goes beyond the examination of the relation between style investing and stock returns under the new perspective of the equilibrium return premium instead of as a deviation from fundamentals. It is also an empirical examination of Merton's prediction. This is the first study that examines the effect of limited inattention and limited participation on stock returns at the institutional level. Taking into consideration that institutional ownership is very high nowadays, the examination of how its characteristics (like style – type categorizations) affect stock returns is very important.

The results of Chapter 2 confirm that higher levels of exposure to stylization lead to higher price volatility and lower liquidity, which is a new result for the literature of style investing. One main contribution of chapter 2 is that instead of assuming that all stocks have the same exposure to a particular style and instead of relying on fund flows and previous style-returns to estimate the effects on stock prices, we examine directly the relation between exposure to styles and volatility-illiquidity.[3] Style investing is a major reason for non-fundamental demand and thus the connection between it and volatility and liquidity is very important for both academics and practitioners.

Chapter 3 of the dissertation departs from the issue of investment styles in the stock market and focuses on the effect of stock liquidity on stock returns, especially during large market declines. The easiness to trade (measured either with the time it takes to find a transaction counterparty or with the monetary cost of trade) is related to stock liquidity. Chapter 3 is essentially based on two branches of literature the one that studies the relationship between stock liquidity and expected stock returns and the other which is part of the limits-to-arbitrage literature and shows the impact of liquidity on stock returns after market shocks.

Amihud and Mendelson (1986) show that less liquidity predicts higher expected returns. In their setting, there are two main differences from a standard asset pricing model; there is cost of trade and there are two types of investors, the short term type and the long term type. The cost of trade acts as a series of negative future cash flows, which are discounted to the present and mechanically reduce the price. The cost of trade is bigger for the short term investors because they incur it more times per period, compared to the long term investors.[4] The level of the stock price depends on the level of the trading costs and the

[3] Wahal and Yavuz (2013) move towards this direction, using the height of the style comovement of each stock as an instrument for the level of its exposition to style investing. Nevertheless, the comovement of each stock to its respective style, crucially depends on the classification of stock to the specific style.

[4] As an example a trader who transacts 10 times during a year pays 10 times higher trading costs from a trader who transacts only 1 time during a year.

portion of the short vs. long term investors on each stock. Long-term investors can amortize these costs in long-horizons and thus gain from the lower prices of illiquid stocks. Short-term investors prefer the liquid stocks because cost of trade could be a large part of their returns. Heterogeneity of the investors (regarding their investment horizons) combined with the cost of trading, produces a limited participation effect which is in essence the same as the one described by Merton. Only long-term investors keep illiquid stocks, while every investor is interested in liquid ones. This clientele effect creates the pricing of stock liquidity.

Stock liquidity is an important characteristic valued by investors, also in a world where arbitrage is limited. The literature of limits-of-arbitrage describes (see Gromb and Vayanos (2010) for a survey), that after a shock in the market, a number of different mechanisms (agency problems – fund redemptions, internal risk controls – automated reduction of positions in stocks, logarithmic utility functions – wealth effects, funding constraints – deleverage, etc) could be activated, forcing investors to sell. Liquidity provides investors the safety to sell their stocks to prices close to their fundamental values. After a market shock occurs, the possibility of forced sales increases, thus investors tilt towards liquid stocks, creating a substitution effect known as "flight-to-liquidity". The spread between the prices of liquid and illiquid stocks widen, the contemporaneous returns of liquid stocks are higher relative to that of illiquid ones and the expected returns of the illiquid stocks become higher relative to that of liquid ones.

Chapter 3 of the dissertation adds new evidence to the earlier literature by emphasizing that a number of investors may be forced to sell stocks immediately after the outburst of a market shock, like the financial crisis of 2008-2009.[5] These investors choose to sell liquid stocks in order to minimize their losses originating from forced selling. In addition, investors who do not face immediate funding needs may choose to exit the stock market for precautionary reasons. In a falling market, they choose to sell mainly liquid stocks to minimize the negative price impact. The prices of liquid stocks may thus fall more than the prices of illiquid stocks due to intense selling, creating an effect in the opposite direction of the "flight-to-liquidity"; let us call it "flight-from-liquidity." This is the first study to provide evidence on the dual role of liquidity during a crisis. It also explains the reason liquidity is a bad predictor of stock returns during crises.

[5] Ben-David et al. (2011), Jotikasthira et al. (2012), Lou and Sadka (2011).

Chapter 3 provides evidence on the simultaneous existence of both effects in the abnormal returns of the stocks during the financial crisis of 2007-2009, and especially after the collapse of Lehman Brothers. We first measure the risk-adjusted returns during the crisis and test if ex ante liquidity can predict returns during crises. The results reject that liquidity is significant for the prediction of stock returns during the crisis. We then measure the abnormal trading activity and its price impact during the crisis and test if ex ante liquidity can predict them. It seems that more liquid stocks exhibit more severe price reduction due to abnormal trading (selling) activity. This is evidence from flight-from-liquidity. We then estimate the risk and turnover adjusted returns during the crisis (as the difference between the two abovementioned parts, namely the risk adjusted returns minus the price impact of abnormal trading) and test if ex ante liquidity can predict them. It seems that more liquid stocks exhibit higher risk- and turnover- adjusted returns during the crisis. This is evidence of flight-to-liquidity.

The results of Chapter 3 also add to the literature relative with the predictive ability of liquidity risk on stock returns. Previous studies define flight-to-liquidity phenomenon as a high correlation between liquidity per se and liquidity risk (Amihud (2002), Acharya and Pedersen (2005)). Their rationale is that liquid stocks carry less liquidity risk, thus they lose less when market liquidity falls. However, their results refer to normal periods, while we focus on a crisis period, and show that liquidity risk is also connected with flight-from-liquidity, exactly due to its correlation with liquidity per se. Investors seek to sell liquid stocks, which also carry less liquidity risk, thus the final outcome show that liquidity risk also predicts flight-from-liquidity.

Finally, previous evidence shows that the liquidity premium is mainly driven from the sell side (Brennan et al. (2012), Brennan et al. (2013)). It is thus an insurance premium regarding the possibility of an immediate sale rather than a symmetric premium related generally with trading (and independent of the direction of the trade). However, the results of Chapter 3 show that liquidity measured by simple market data during normal periods is not an adequate insurance for the investors during large market declines. Liquid stocks also fall during large market declines due to the concentrated sales of their owners. The need of more sophisticated measures which take into account the funding needs of their owners should be used for better prediction of stock returns during crises.

The three chapters of the dissertation provide new evidence on the behavior of asset prices. Ownership, liquidity and the inter-connection between classical asset pricing and limits-of-arbitrage drive the results. These issues seem at the core of the academic landscape over the next few years, providing impressive new explanations for the stylized facts in the markets.

Chapter 1: Style Concentration of Ownership and 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.[6] 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.

[6] 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.

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.[7]

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.[8]

---

[7] 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

[8] 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.

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.[9] The 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.[10] 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

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

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

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 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.[11]

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 maximum of 1.00) at the value 0.50. The results remain similar, although now the coefficient of style concentration is higher.[12] 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.

[11] 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).

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

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.

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.[13] 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.[14]

---

[13] 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.

[14] 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

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.[15] 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

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.

[15] 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.

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 matter, would deliver the same concentration parameter of 1/5 = 0.2 in both cases and would miss a lot of the information.[16]

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.[17] 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

---

[16] 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.

[17] 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.

Index as a measure of the concentration of ownership and study its relation with stock liquidity.

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.[18] 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[19] and are based on the mandatory 13F filings.[20] 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.[21] 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

---

[18] 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.

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

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

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

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 by those who file, based on their portfolio characteristics.[22] The data base uses thirty two different style options for the classification of institutional investors.[23]

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*[24] *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." [25]

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

---

22 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.

23 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.

24 First Call is a Thomson First Call is a branch of Thomson Financial and it is a major provider of estimates.

25 http://www.tfsd.com/marketing/banker_r2/HomeFAQs.asp

classification of a fund. However, they find that the approach which is based on the portfolio characteristics, predict fund returns better.

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.[26]

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

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

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 dollar volume (which is the shares volume times the price of the stock).[27] 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.[28]

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).[29]

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

---

[27] $ILLIQ_{i,q} = 1/D \cdot \sum_{d=1}^{D} |r_{i,d}| / \$volume_{i,d} \cdot 10^6$, 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.

[28] 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: $standILLIQ_{i,q} = \frac{ILLIQ_{i,q} - \overline{ILLIQ_q}}{s.d.(ILLIQ)_q}$.

[29] 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.

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.

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: $ILLIQ_{i,q} = 1/D \cdot \sum_{d=1}^{D} \lvert r_{i,d} \rvert / \$volume_{i,d} \cdot 10^6$, 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.

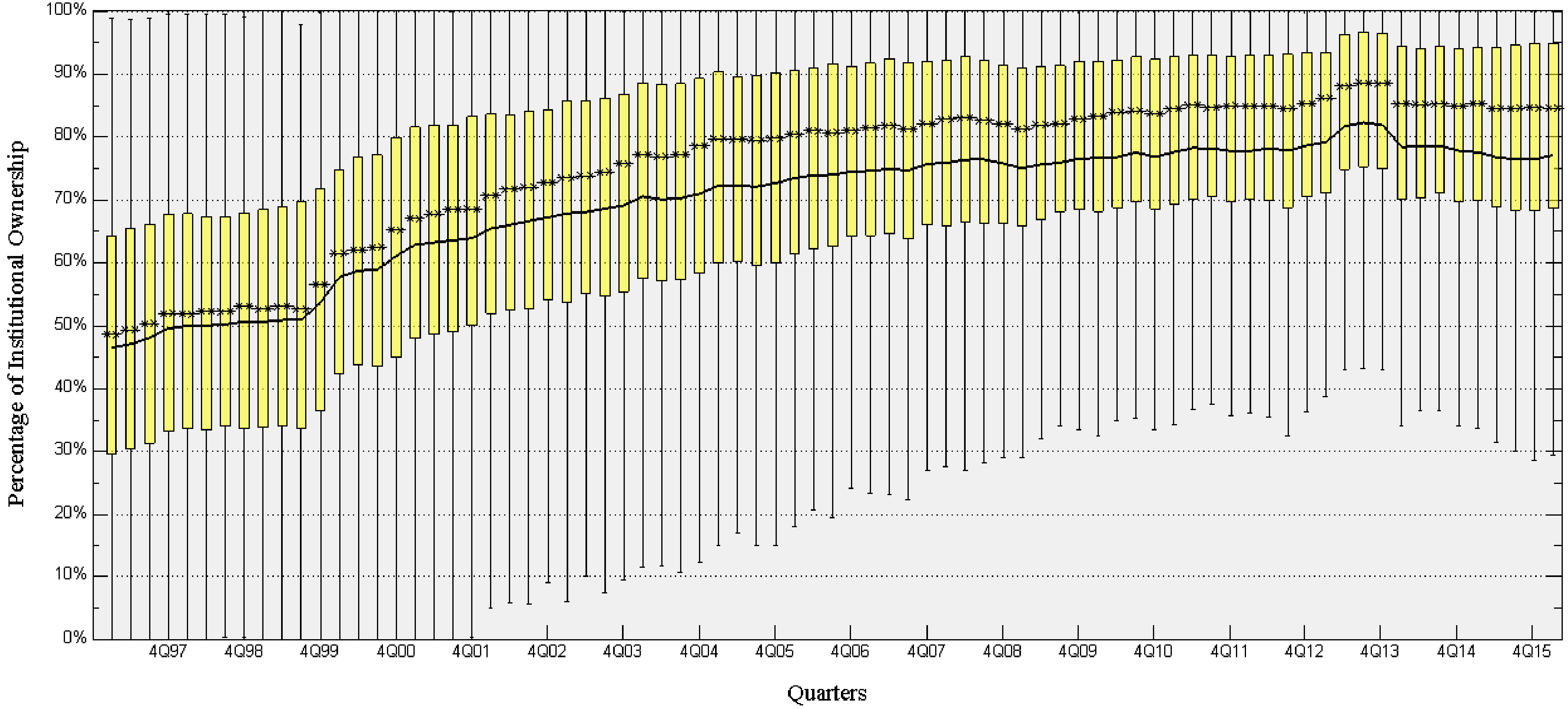
Figure 4 traces the cross-sectional distribution of H over time. Mean concentration was gradually reduced from around 0.29 in the early years to slightly above 0.21 today. This is a substantial reduction in market-wide concentration, indicating that over the years, stocks are chosen by a more diversified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

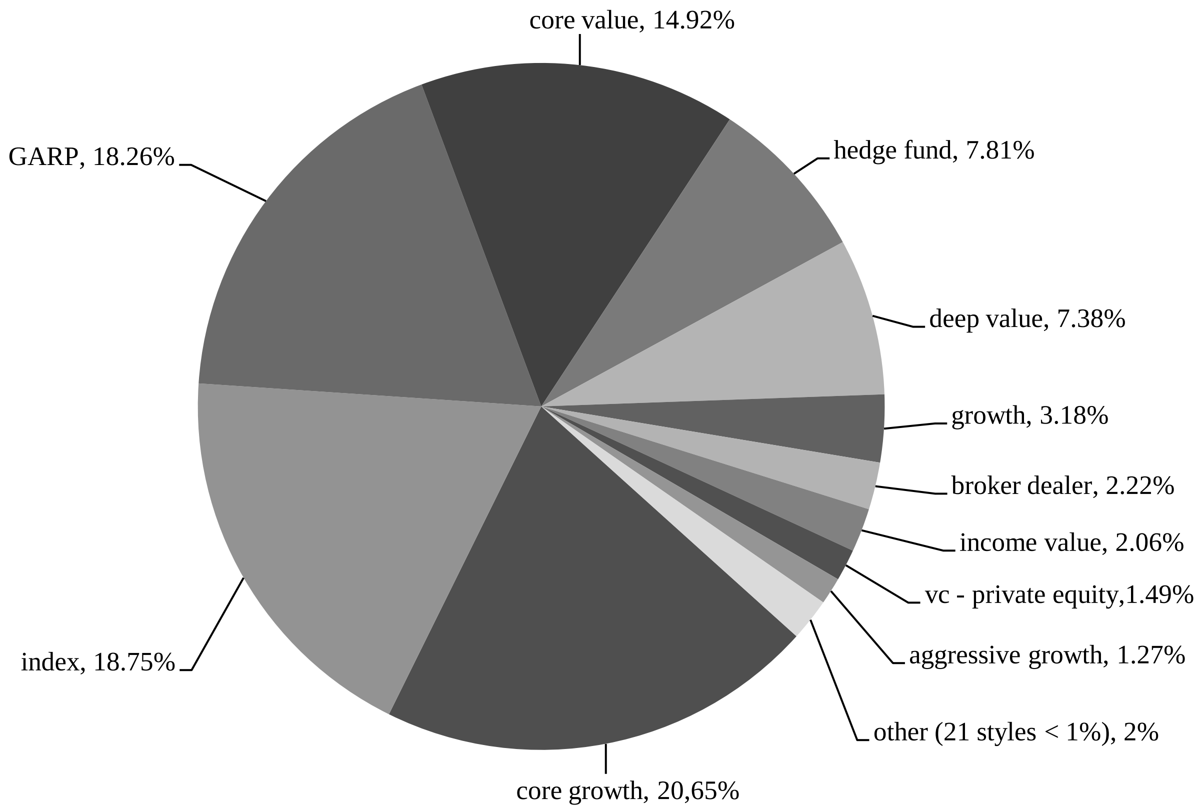
Table 5 contains the correlations of H with the stock ownership percentages of the large investment styles. As shown, H is not significantly correlated with any individual investment style. Its highest correlation is with the ownership of the Index style. This correlation is negative, at - 0.23. Apparently, 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



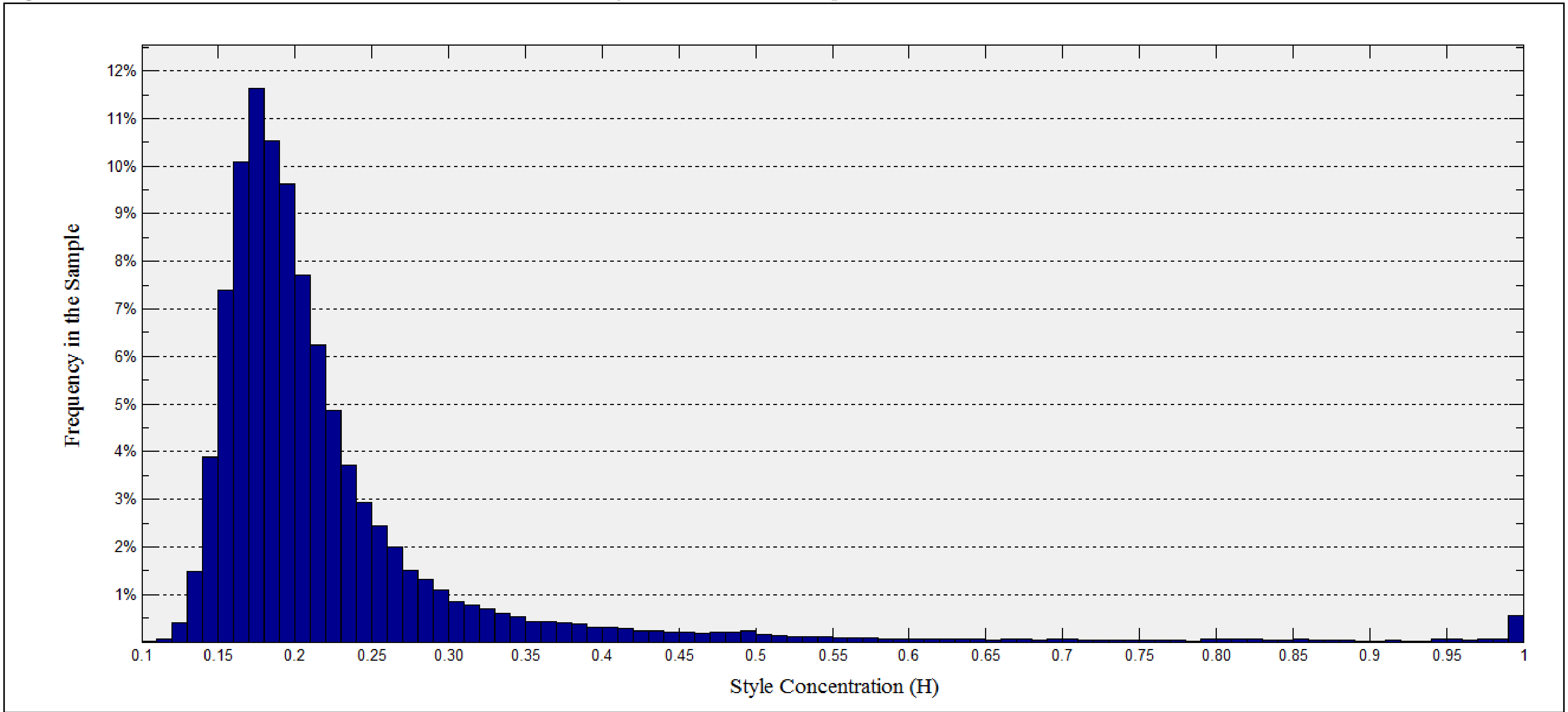
The figure illustrates the evolution of the distribution of institutional ownership over the 77 quarters of the sample (1997-Q1 to 2016-Q1). The solid black line represents the cross-sectional mean of institutional ownership for each quarter. Black stars represent the median institutional ownership in 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.



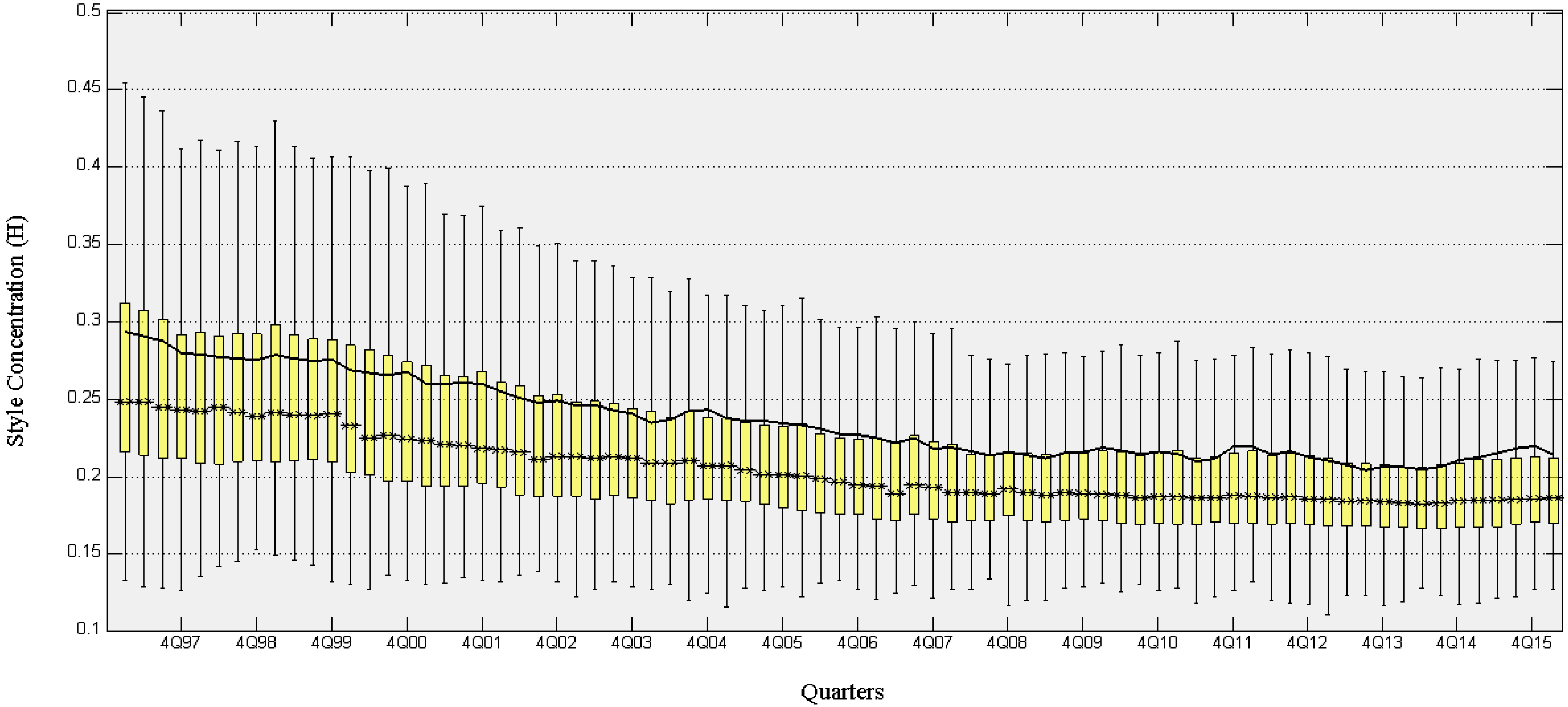
The figure illustrates the mean percentage shares of the investment styles in the pooled sample. Percentages add to 100%. . The sample covers 77 quarters, from the 1997-Q1 to the 2016-Q1 and contains 72,880 observations of stocks (an average of 946 stocks per quarter).

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 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.[30] 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.[31] 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

[30] 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.

[31] 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.[32] 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.

[32] 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.[33]

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.[34] 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.[35] 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

[33] 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.

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

[35] 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.[36] 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

[36] 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.[37]

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.[38] 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.[39] 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.[40]

---

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

[38] 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.

[39] 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).

[40] 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 $\sum_{j=1}^{K}\left(x_j - \bar{x}\right) = 0$.

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:

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

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 $\frac{1}{MV^2} \cdot \sum_{j=1}^{K}\left(x_j - \bar{x}\right)^2 = 0$ 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	45,355

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).[41]
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,

[41] 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.[42] 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

[42] 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, B. H. (2011), Style-Related Comovement: Fundamentals or Labels? The Journal of Finance 66, 307–332.

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.

Chapter 2: Style Concentration in Stock Ownership, Stock Price Volatility and Liquidity

1. Introduction

Limited participation in the stock market may give rise to stock illiquidity and stock price volatility, due to an imperfect absorption of possible excess demand for stocks, when stock owners are hit by liquidity shocks. A large number of theoretical and empirical studies show that, indeed, often the demand slopes for stocks are not flat, as it is assumed by the classical finance view.[43] It is thus of great importance to understand the determinants of limited participation and how they affect the liquidity and volatility of stocks.

Over the last 30 years, institutional ownership of stocks has risen significantly. In our sample, the average institutional ownership increased from around 45% in 1997 to almost 78% in 2015.[44] A major characteristic of institutional ownership is the self-definition of the institutions into different investment styles. The investment styles are distinguished by their legal status or their trading strategy, i.e., hedge funds, venture capital funds, broker-dealer funds, index funds, growth or value funds, big-stock, mid-cap-stock or small-stock funds, etc. In many cases, the funds carry in their names the style they follow.

The widespread adoption of investment styles by institutional investors creates a style-related inattention. Stocks that are followed by a larger number of different investment styles receive more investment attention relative to stocks that are followed by only few investment styles. Consequently, there is less capital available for the stocks that receive limited attention and thus they become less liquid and their prices are more volatile.

Style investing is the focus of a long academic literature, which reports evidence of herd trading, momentum, reversals and large co-movement among stocks belonging to the

[43] Shleifer (1986), Harris and Gurel (1986) show that the inclusion or the exclusion of a stock in a stock index induce a price increase or decrease, respectively. Coval and Stafford (2007), Frazzini and Lamont (2008) and Lou (2012) show that the flow driven demand for stocks by mutual funds induce a price impact. Greenwood and Vayanos (2009) show that the demand for bonds also induces a price impact.

[44] 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.

same style ((Barberis and Shleifer, 2003) and others). This literature has not examined the issue of stock inattention generated by the existence of investment styles. Yet, stocks that are owned by few styles are more exposed to the price pressure of style investing, since there are only a few owners to absorb the possible excess demand for the stock. Inattention is high for these stocks. At the other end, stocks with dispersed style ownership are less exposed to the price pressures of style investing, since the demand of the different styles offset each other, reducing the final excess demand.

In this paper we examine the relation between style inattention (measured by the style concentration in stock ownership) and the volatility of stock prices and the stock liquidity (measured either as the cost of trade or as the trading activity). We measure style concentration in stock ownership (henceforth H) as the Herfindahl Index of the shares of each style in each stock. Our hypothesis is that a higher H is related with a higher stock price volatility (measured as the daily volatility within each quarter), less trading activity and higher cost of trading (measured with two different proxies, the ILLIQ (Amihud, 2002) and the bid-ask spread as percentage of the stock price).

The rationale of our paper is that the price of a stock with high style concentration in its ownership is more sensitive to the flows of style investing, since there is insufficient number of alternative styles to absorb the excess demand. On the other hand, the price of a stock with its ownership being dispersed among a large number of different styles is less sensitive, since the excess demand in this case is smaller. In addition, different styles may even have negatively correlated flows, which in turn reduce even more the excess demand.

Our econometric analysis is based on pooled-OLS regressions of the above three dependent variables in quarter (t+1) on the inattention variable H and other control variables of quarter (t). Our results indicate a statistically and economically significant relation between style concentration in ownership and stock price volatility (either the total or the idiosyncratic one). For example, in the univariate case, one standard deviation change in H predicts 19.3 b.p. higher future daily idiosyncratic volatility, with a t-statistic of 6.11. Similarly it predicts a rise in price bid-ask spread by 0.3 standard deviations, with t-statistic of 15.98.

Our results are in line with previous empirical work, which explores the relation between ownership and either volatility or liquidity. Greenwood and Thesmar (2011) construct a measure that is based on the volatility of the fund flows of the owners of a stock, weighted by the correlation between the fund flows of the owners and the ownership concentration.[45]. They find that a higher value of their measure is connected with higher future stock price volatility. Amihud et al. (1999) find that a decrease in ownership concentration results not only on lower expected returns but in higher liquidity as well. Rubin (2007) finds that an increased level of institutional ownership increases liquidity but increased ownership concentration decreases liquidity. Our paper complements this strand of literature by showing that the stylization of stock ownership (which creates a form of ownership concentration) also affects volatility and liquidity.

As mentioned earlier, our study is also related with the strand of the literature which examines the role of style investing on financial markets. Barberis and Shleifer (2003), in their novel paper, show that the existence of style investors in a market creates style momentum, style reversal and excess comovement among the stocks of the same styles. Subsequent empirical studies confirm their findings (Teo and Woo (2004), Froot and Teo (2008), Boyer (2011), and Wahal and Yavuz (2013)). The present study extends the earlier work by examining the relation of style investing with the volatility and liquidity of stocks.

Finally, our paper also contributes to the literature that examines the relation between institutional investing and stock price volatility (Sias (1996), Bushee and Noe (2000), Gabaix et al. (2006)), showing that not only the level of institutional ownership matters, but also its structure across different styles and types of investors. The paper is also related to the paper of Allen and Gale (1994), who show that limited participation in a market could result in increased volatility of asset prices, and to the paper of Mitchell et al. (2007), who argue that slow moving capital (from asset to asset) could result in higher price impacts, thus higher price volatility and illiquidity.

The remainder of the paper is organized as follows: Section 2 describes the construction of our variable of style concentration. Section 3 describes our data and the variables that we use. Section 4 provides a preliminary analysis of our variables and illustrates some stylized facts about style investing. Section 5 presents the main econometric results about the relation between style concentration and stock price volatility. Section 6

[45] They name this variable "stock price fragility". If the liquidity shocks of the owners of a stock are volatile, given that these shocks induce a price impact, the price of this stock will be volatile too. In addition, the structure of the ownership of the stock determines to which extend the volatility of the liquidity shocks passes to the volatility of the price. A small number of owners or a high correlation of their liquidity shocks will amplify the effect of the liquidity volatility of the owners to the stock price volatility.

presents the main econometric results on the relation between style concentration and stock liquidity. Finally, Section 7 discusses our findings and concludes.

2. Style Concentration

The classical approach of style investing studies is to classify each stock to a sole style according to its characteristics. However, in reality, stocks are held by funds of different styles at the same time, thus they are exposed to a continuum of styles. The classification of stocks into styles neglects this information and implicitly assumes that each stock is maximally exposed to style investing. Instead of classifying each stock to a unique style, we utilize the characterization of the stock owners into styles and subsequently measure the exposure of each stock to the "stylization" of the stock market. We are thus able to detect cross-sectional differences on the exposure of the stocks to stylization, something that it is not possible with the traditional approach, which implicitly treats all the stocks as having the same (maximum) exposure.

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.[46] 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.[47]

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 that 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 exposure of a stock to the segmentation that is created due to the existence of different styles and strategies by the market participants. Theoretically, a stock which is held by all the market participants, it has the minimum H value and it is effectively immune to the stylization. On the opposite extreme a stock that is held only by owners that follow a sole style, it has the maximum H value and it is fully exposed to the stylization of the market.

The exposure to the market stylization is connected with two different (although relative) effects. The first effect is the direct relation between the readily available capital for a stock and its liquidity/price volatility. If a stock is highly exposed to the stylization, less capital is readily available to meet the excess demand, because only a part of the market participants follow the stock, which in turn decreases the liquidity and increases the non-

---

[46] Leaving the smaller investors out of the calculation of index *H,* makes the implicit assumption that the excluded investors do not cause changes in the ownership 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.

[47] 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.

fundamental price volatility of the stock. The second effect has to do with the nature of style investing per se. It is known that there is a continuum of styles, which are defined in multiple dimensions, resulting in a rich correlation structure between their demands for stocks. The excess style-related demand for a stock that is held by every different style is very low, since at the same time the buying needs for the one style are offset by the selling needs of another style. As a result, the liquidity of this stock further increases and the non-fundamental volatility of this stock is further reduced. The two effects move towards the same direction and are the base of our hypothesis, that H predicts higher stock price volatility and lower stock liquidity.

The use of the 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. Our study is the first to use the Herfindahl Index in a higher level of aggregation in order to measure the effects of ownership segmentation to the risk and liquidity of stocks.

3. 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.[48] The sample consists of 1,295 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

[48] The maximum number of quarters used in the panel analysis is 76 and not 77, as volatility and illiquidity variables 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.

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 805 stocks as some of the independent variables are missing.

3.1 Institutional Data

Data for institutional investors are from Thomson Reuters[49] and are based on the mandatory 13F filings.[50] 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.[51] 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 by those who file, based on their portfolio characteristics.[52] The data base uses thirty two different style options for the classification of institutional investors.[53]

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*

---

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

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

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

[52] 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.

[53] 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.

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*[54] *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." [55]

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.

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 fractions 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.

3.2 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.[56]

---

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

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

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

Total volatility is estimated as the standard deviation of the daily stock returns within each quarter. 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.

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 dollar volume (which is the shares volume times the price of the stock).[57] ILLIQ is an illiquidity measure[58] 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.[59]

The bid-ask spread is estimated as the quarterly mean of the daily difference between the end of the day ask and bid price, divided by the end of the day price of the stock. The bid and the ask prices are from Bloomberg. We divide the spread by the price in order to measure it as a percentage of the price, excluding the effect of the level of the price from its value. Spread also decreases considerably after the first years of our sample, thus we also cross-sectionally normalize its value for each quarter.[60]

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

---

57 $ILLIQ_{i,q} = 1/D \cdot \sum_{d=1}^{D} |r_{i,d}| / \$volume_{i,d} \cdot 10^6$, 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.

58 A higher value of ILLIQ indicates lower stock liquidity.

59 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: $standILLIQ_{i,q} = \frac{ILLIQ_{i,q} - \overline{ILLIQ_q}}{s.d.(ILLIQ)_q}$.

60 We estimate the normalized bid-ask spread for each quarter by subtracting the cross-sectional mean of bid-ask spread of that quarter and then by dividing with the cross-sectional standard deviation of that quarter: $standSpread_{i,q} = \frac{spread_{i,q} - \overline{spread_q}}{s.d.(spread)_q}$.

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.

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 liquidity 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).[61]

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.

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.

[61] 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.

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
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
Total Volatility (tot_vol)	The standard deviation of the daily stock returns within quarter q.	Prices from Bloomberg. (Bloomberg Datatype: PX_LAST)	79,698
Idiosyncratic Volatility (idio_vol)	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
Share Turnover (turnover)	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: $turnover_{i,q} = 1/D \cdot \sum_{d=1}^{D} volume_{i,d}/(total\ \#\ of\ shares)_{i,d}$, where D is the total number of trading days during the quarter q.	We take the trading volume and the total number of shares from Bloomberg. (Bloomberg Datatypes: PX_VOLUME and EQY_SH_OUT, respectively)	78,623
ln(turnover)	The natural logarithm of the share turnover (turnover).	We take the trading volume and the total number of shares from Bloomberg. (Bloomberg Datatypes: PX_VOLUME and EQY_SH_OUT, respectively)	78,623
ILLIQ (Amihud,2002) (ln(ILLIQ))	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	Stock prices from Bloomberg. (Bloomberg Datatype: PX_LAST) Share volumes from Bloomberg.	79,719

	scaling factor of 10^6: $ILLIQ_{i,q} = 1/D \cdot \sum_{d=1}^{D} \lvert r_{i,d} \rvert / \$volume_{i,d} \cdot 10^6$, where D is the total number of trading days during the quarter q.	(Bloomberg Datatype: PX_VOLUME)	
Bid-ask spread (ln(spread))	The natural logarithm of the Bid-Ask Spread. Spread of stock i for quarter q is the average of the daily spread, as a percentage of stock price.	Stock prices, bid prices and ask prices from Bloomberg.	76,289
Size (ln(mv))	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))	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 (momentum)	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))	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
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	MOM data directly from the site of Kenneth French: http://mba.tuck.dartmouth.edu/pages/faculty/	255

	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).	ken.french/data_library.html#Research	
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 (illiq-beta)	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

4. 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.

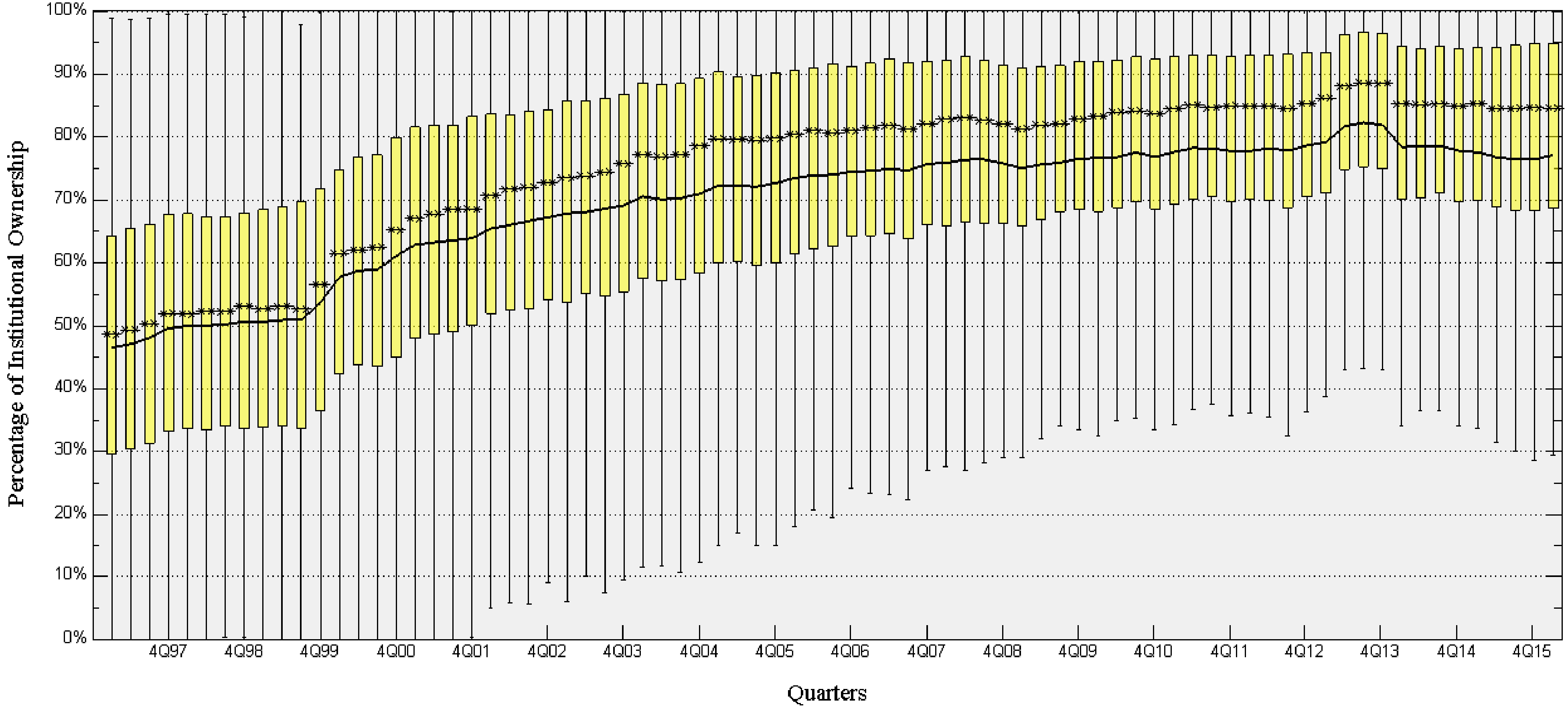
Figure 4 traces the cross-sectional distribution of H over time. Mean concentration was gradually reduced from around 0.29 in the early years to slightly above 0.21 today. This is a substantial reduction in market-wide concentration, indicating that over the years, stocks are chosen by a more diversified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

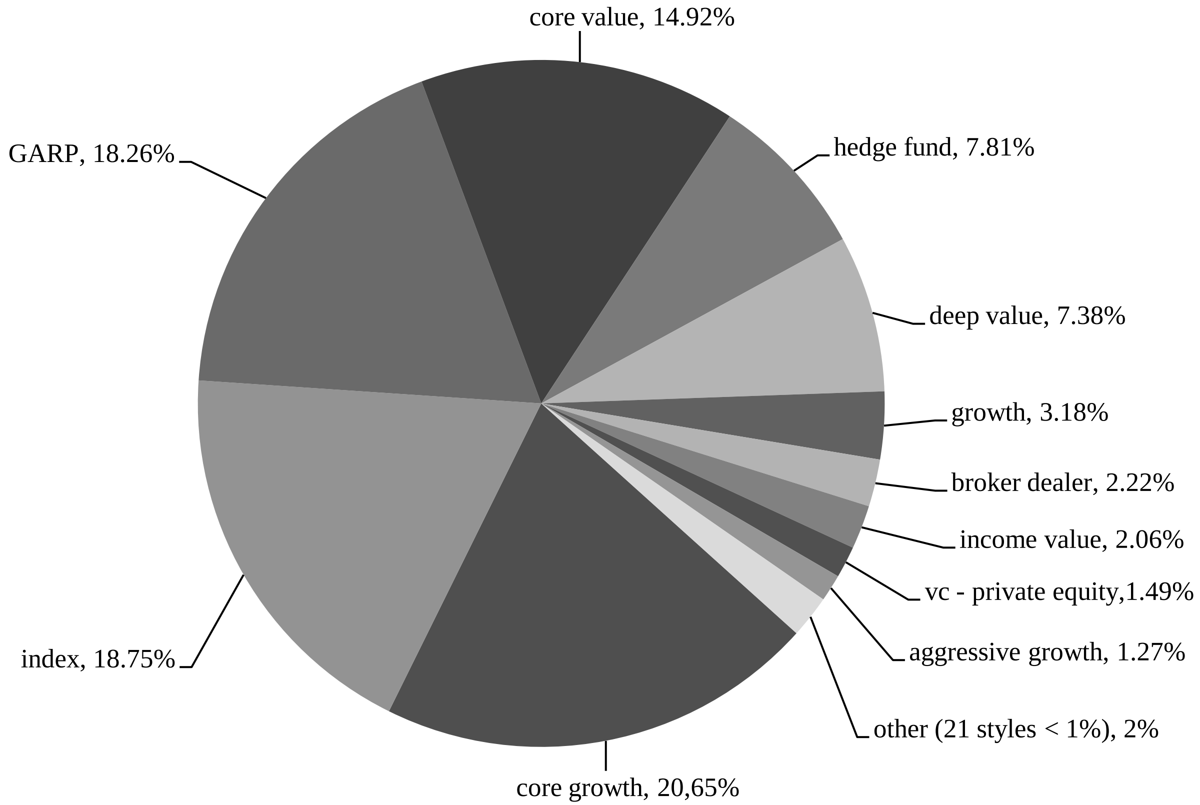
Table 5 contains the correlations of *H* with the stock ownership percentages of the large investment styles. As shown, *H* is not significantly correlated with any individual investment style. Its highest correlation is with the ownership of the Index style. This correlation is negative, at - 0.23. 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



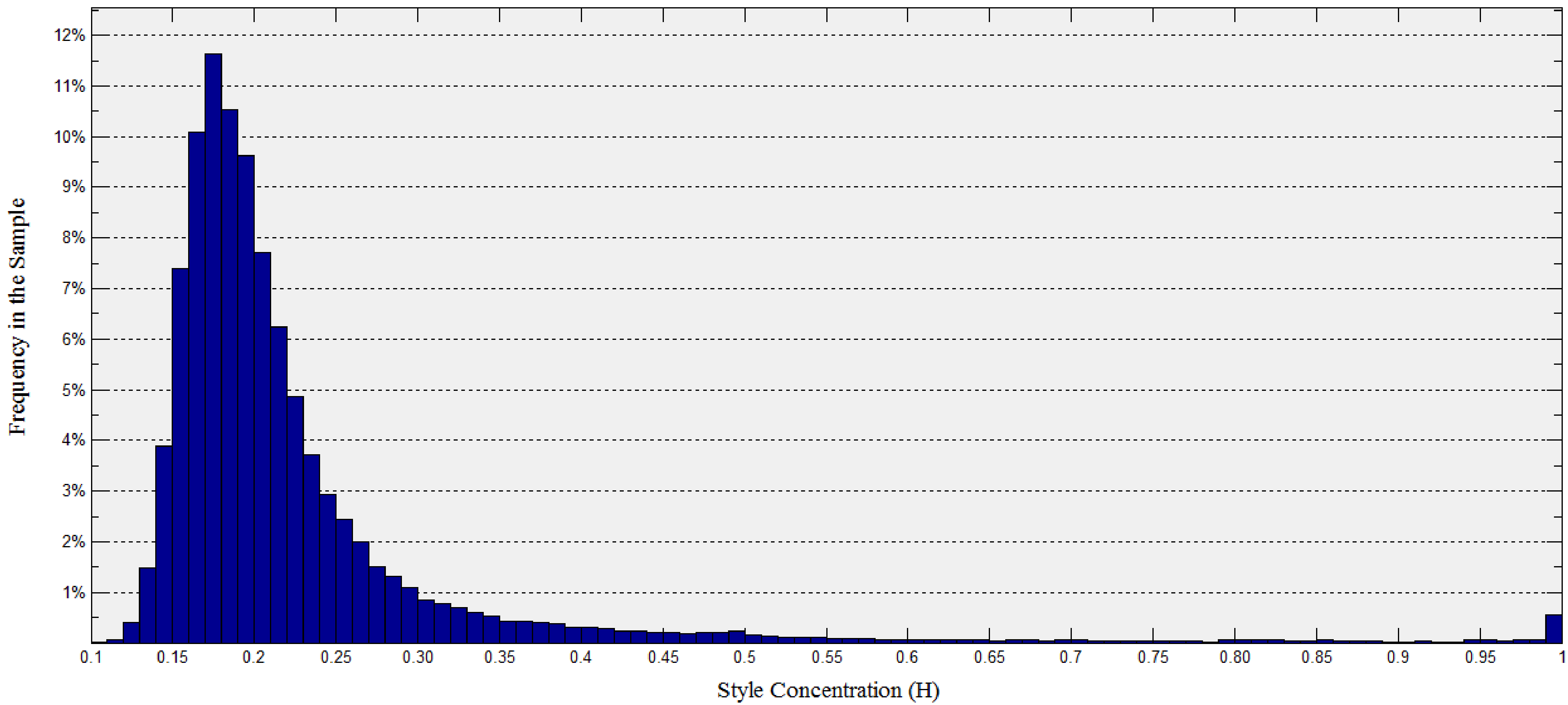
The figure illustrates the evolution of the distribution of institutional ownership over the 77 quarters of the sample (1997-Q1 to 2016-Q1). The solid black line represents the cross-sectional mean of institutional ownership for each quarter. Black stars represent the median institutional ownership in 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.



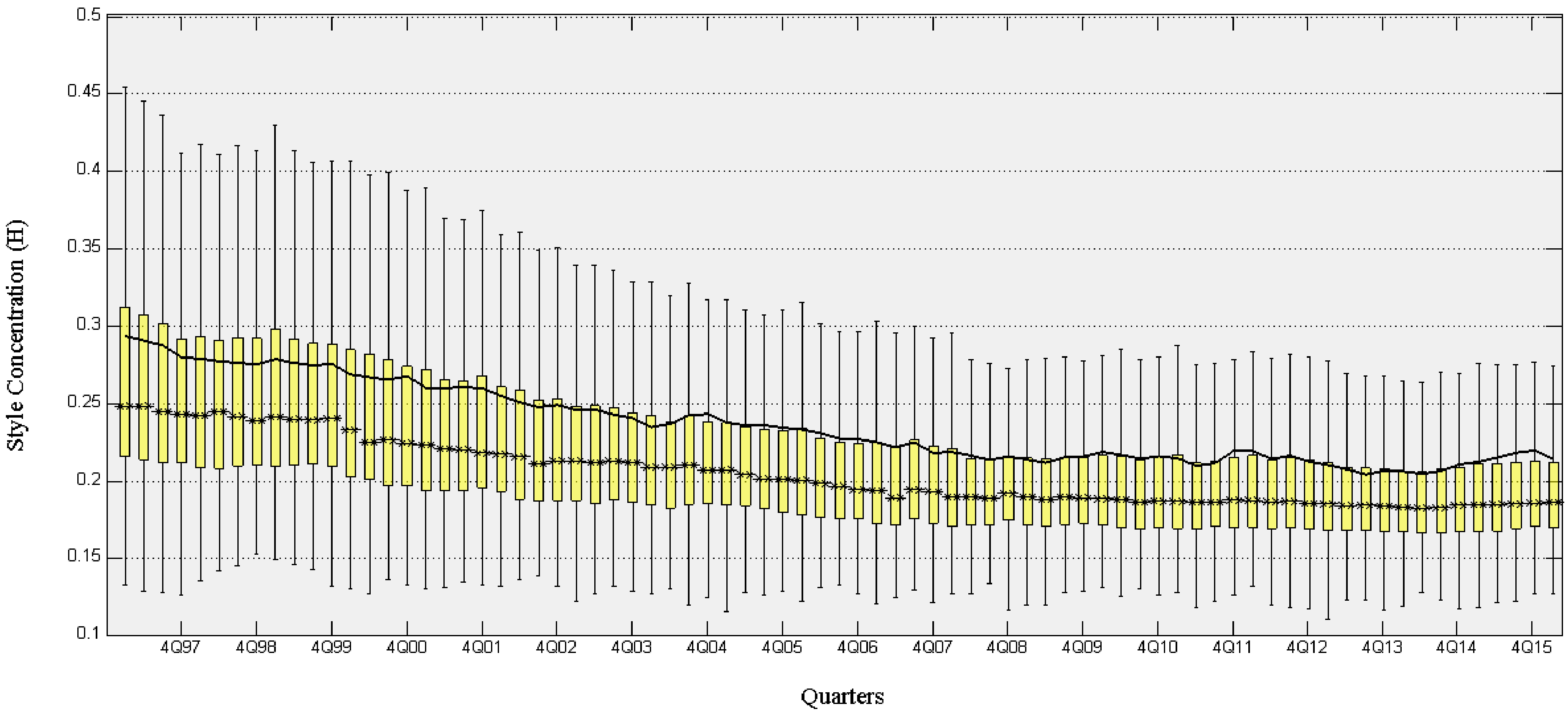
The figure illustrates the mean percentage shares of the investment styles in the pooled sample. Percentages add to 100%. . The sample covers 77 quarters, from the 1997-Q1 to the 2016-Q1 and contains 72,880 observations of stocks (an average of 946 stocks per quarter).

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 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*
H	0.233	0.118	3.949	21.774	0.111	0.199	1.000
tot_vol (%)	2.411	1.560	3.647	34.584	0.010	2.030	38.380
idio_vol (%)	2.057	1.381	4.119	45.974	0.030	1.720	36.760
ln(ILLIQ)	~0.000	0.999	0.726	3.919	-2.529	-0.108	6.544
ln(spread)	~0.000	0.999	0.377	5.360	-6.028	-0.080	6.267
turnover (%)	0.791	1.893	86.286	9604.244	0.000	0.580	243.920
ln(turnover)	-5.306	1.129	-1.768	10.114	-14.044	-5.157	0.891
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
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
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 805 stocks per quarter.

	H	tot_vol	idio_vol	ln(ILLIQ)	ln(spread)	turnover	ln (turnover)	ln(mv)	ln(mtb)	market beta	SMB beta	HML beta	MOM beta	illiquidity beta	mom/tum	ln (dta)
H	1															
tot_vol	0.085	1														
idio_vol	0.152	0.953	1													
ln(ILLIQ)	0.430	0.259	0.320	1												
ln(spread)	0.269	0.260	0.313	0.623	1											
turnover	-0.075	0.146	0.125	-0.100	-0.042	1										
ln (turnover)	-0.480	0.222	0.155	-0.488	-0.275	0.342	1									
ln(mv)	-0.293	-0.362	-0.422	-0.868	-0.535	-0.005	0.164	1								
ln(mtb)	-0.058	-0.189	-0.163	-0.309	-0.264	-0.065	0.020	0.364	1							
market beta	-0.065	0.163	0.131	-0.028	0.019	0.062	0.144	-0.012	-0.065	1						
SMB beta	0.026	0.149	0.150	0.253	0.167	0.037	0.048	-0.287	-0.067	-0.134	1					
HML beta	0.019	0.067	0.080	0.098	0.090	-0.001	-0.046	-0.122	-0.148	0.080	0.052	1				
MOM beta	-0.011	-0.114	-0.110	-0.060	-0.086	-0.025	-0.018	0.072	0.090	0.072	-0.132	0.080	1			
illiquidity beta	0.002	0.062	-0.076	-0.018	-0.027	-0.204	-0.043	0.026	0.011	-0.162	-0.015	-0.017	-0.004	1		
mom/tum	0.001	-0.097	-0.077	-0.038	-0.079	0.007	0.023	0.041	0.218	0.014	0.025	0.004	0.033	-0.016	1	
ln(dta)	-0.030	-0.062	-0.074	-0.138	-0.056	-0.005	-0.016	0.153	-0.082	0.006	-0.081	0.050	-0.025	0.001	-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

5. Econometric results: stock price volatility and style concentration.

5.1 Equation specification and control variables

We proceed with the analysis of the relation between H and the stock price volatility. Since our goal is to study whether the style concentration has predictive ability over the volatility we regress the idiosyncratic volatility of quarter $q+1$ on the style concentration of quarter q:

$$\sigma_{i,q+1} = \alpha + b \cdot H_{i,q} + e_{i,q+1} \quad (3)$$

We expect that the style concentration is positively correlated with the volatility of the stock price, thus our hypothesis is that:

$$H_0 : \beta > 0 \quad (4)$$

In addition we test a number of different specifications, in which we include a number of control variables (Z). These variables are either known determinants of stock price volatility, or they are related with the investment styles, or both. We also include the own lag value of volatility, to address the endogeneity concerns, especially since volatility is highly persistent:

$$\sigma_{i,q+1} = \alpha + b \cdot (H)_{i,q} + \Lambda' \cdot Z_{i,q} + e_{i,q+1} \quad (5)$$

The natural logarithm of market capitalization is included as a control variable, since it is known that the prices of small stocks tend to be more volatile compared to the prices of large stocks. In addition, size is one of the major determinants of investment styles, thus its inclusion controls for any effect derived from specific styles. Moreover, we include the natural logarithm of market-to-book ratio, which is also a basic determinant of investment styles, to further control for the effects of specific styles.

The systematic risk is a major determinant to the stock price volatility, thus we also control for this, with the inclusion of the market beta to our specifications. In the complete specification case we also include the rest of the betas of a four-factor model (Carhart, 1997) to capture additional aspects of the effect of systematic risk on the stock price volatility.

The price of the relatively illiquid stocks exhibit higher price volatility, since any transaction on them induces a higher price impact. We thus include either the ILLIQ (the price impact measure of Amihud) or the bid-ask spread in order to control for the effect of liquidity on stock price volatility. In addition, we also include the illiquidity beta of the stocks to control for the effect of liquidity risk in addition to the effect of liquidity per se. We furthermore include the natural logarithm of the share turnover of each stock to control for the effect of the trading activity. Higher trading activity is related to higher price volatility, either because of the frequent arrival of new information or just from the frequent liquidity trading.

The natural logarithm of the debt-to-assets ratio is also included to some of the specifications to capture the leverage of the firms which also may affect the volatility of the prices of their stocks. We also include the price momentum to the full specification case in order to control for the attention that is created by past stock returns and due to its use as a key variable for the determination of some styles.

To further control for the possibility that certain investment styles are related with the stock price volatility, we include to some of the specifications the ownership shares of each investment style. This is a very strict control, since the H is constructed by these shares. However, with the inclusion of them into the specification, we ensure that the coefficient of H shows the effect of style concentration and it is not affected by any specific style.

We furthermore include the lagged value of the stock price volatility, since it is a highly persistent variable. With the inclusion of the lagged volatility we also address to some extend the possible endogeneity. It is possible that stocks with high price volatility are riskier and thus only investors with high quality of information hold them. Since we analyze the ownership structure in the style level, this is not a severe problem, but we could not exclude the case that volatile stocks attract the attention of only few styles, which would result in high style concentration. In the absence of a meaningful control, the inclusion of the lagged value of volatility gives us confidence that the effect of H is not driven by the volatility itself.

5.2 Main econometric results

We run pooled time series – cross sectional OLS regressions, including 75 quarterly dummies in order to address the time effect.[62] As a consequence of the time effect, the observed Adjusted R^2s are very high, even if we do not include the lagged value of the volatility (in the results of Table 6, they range from 21.4% to 68.6%). The inclusion of the lagged value of volatility increases even more the Adjusted R^2s, since volatility is highly persistent.

The stock price volatility is serially correlated from quarter to quarter, thus we address the presence of the firm effect (Cochrane (2001), Petersen (2009)) by estimating t-statistics

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

that are based on robust (heteroskedasticity-consistent) standard errors, clustered at the firm level. This correction addresses both the heteroskedasticity and the serial correlation, which are present in our data, and corrects (reduces) the size of the reported t-statistics. For easiness of exposition, we use three asterisks (***) to denote statistical significance at the 1% level, two asterisks (**) at the 5% level, and one asterisk (*) at the 10% level.

The results (Table 6) confirm our hypothesis that higher style concentration predicts higher stock price volatility (both total and idiosyncratic). Independently of the specification that we use, H enters the regressions with a positive and statistical significant coefficient (at 1% level). Our results are also economically significant since one standard deviation of H predicts from 13.5 b.p. (univariate case, column (1) of Table 6) to 18.7 b.p. (case with main controls, column (3) of Table 6) higher daily total volatility. The results for the idiosyncratic volatility are even stronger with one standard deviation of H predicting from 19.3 b.p. (univariate case, column (6) of Table 6) to 19.5 b.p. (case with main controls, column (8) of Table 6) higher daily idiosyncratic volatility.

When we include the lagged value of volatility in our specifications (columns (2), (4)-(5), (7) and (9)-(10) of Table 6), the coefficient of H remains statistically significant, although smaller in magnitude. However, taking into consideration that in the presence of the lagged value of volatility the coefficient of H should be interpreted as the effect on volatility innovations, the results remain economically significant. Furthermore, the strength of the coefficient of H even in the presence of the lagged value of volatility is evidence that there is not any severe endogeneity problem.

The results for the rest of the variables are as expected and are in line with the evidence of previous studies. Higher market beta, bid-ask spread and turnover predict higher stock price volatility. On the other hand, bigger size and market-to-book ratio predict lower stock price volatility.[63] When we include in our specification the shares of each style (columns (5) and (10) of Table 6) the coefficient of H remains at similar levels, confirming that the result is not driven by the presence of any specific style.[64]

Our results are in line with the results of Greenwood and Thesmar (2011). Style concentration does not include the volatility and the correlation structure of the fund flows, but instead takes into account that the funds of the same style have correlated flows. Thus, one could interpret style concentration as an alternative stock price fragility measure, which

[63] The results for the rest of the control variables are presented in the Appendix and are also the expected ones.

[64] This is a very strict control, since the shares of the styles are the inputs for the estimation of style concentration.

focuses on the effect of style investing. The strong coefficient of H (both in statistical and economical terms), even if H does not include the volatility of the fund flows of quarter q, is evidence that style investing could be one of the major reasons for stock price fragility.

Greenwood and Thesmar also estimate a coefficient for the ownership concentration on individual fund level. This coefficient has a negative sign and very small magnitude in every specification in which it is included. One reason for the striking difference in relation to the coefficient of our H could be that their main variable of stock price fragility ("G") already includes the ownership concentration on individual level. A second reason is that the nature of ownership concentration on individual level is much different from the nature of H. The simple concentration provides very little information about the ownership structure of a stock because the vast majority of stocks are owned by a large number of different funds, each of them holding a small part of the shares. On the other hand, style concentration provides much more information, since there are large cross-sectional differences in the participation of various styles in stock ownership. Additionally, style concentration provides more information about the exposure of stocks to style investing, something that a simple concentration measure does not provide.

Our results also reveal an interesting asymmetry between the effect of H on total and idiosyncratic volatility. One could expect that the effect would be stronger in the first case, since total volatility is higher than the idiosyncratic part, due to the inclusion of the systematic risk. However, we find that the effect of H is stronger in the case of the idiosyncratic volatility. The answer for this asymmetry lies in the relation of H with the exposure of a stock to the systematic risk. The correlation between H and the market beta is negative, which means that stocks with higher H have smaller exposure on systematic risk and vice versa. A different way of looking into this asymmetry is that the proportion of idiosyncratic volatility to the total one, is higher for stocks with higher H.

To our knowledge this is the first paper that studies how the existence of investment styles affects the stock price volatility. Papers on style investing examine how it relates to momentum and reversal effects, as well as how it affects the comovement within the styles. Our study is complementary to this literature, showing that style investing is also related to stock price volatility.

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

Panel OLS regressions of the daily price volatility of stock i at quarter q+1, $\sigma_{i,q+1}$, 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:

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

There are 10 regressions in columns 1 through 10. Columns (1) to (5) show the results for the total volatility and columns (6) to (10) show the results for the idiosyncratic volatility (the residual of a 4-factor model (Fama-French, Carhart)). A time effect with quarterly dummies is included in every regression. The variables of each regression are described in the very left column. Lag-volatility is the lagged total volatility for columns (1) to (5) and the lagged idiosyncratic volatility for columns (6) to (10). 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: betas of SMB, HML and MOM factors, illiquidity beta, price momentum and ln(dta)). See Table 1 for the detailed definitions of the variables.

Volatility is measured in percentage form. The sample covers the period between 1997-Q1 and 2015-Q4 (76 quarters) and on average consists of around 805 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 robust standard errors that are clustered by firm. 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)
	tot_vol	tot_vol	tot_vol	tot_vol	tot_vol	idio_vol	idio_vol	idio_vol	idio_vol	idio_vol
H	**1.15***** **(4.39)**	**0.40***** **(4.48)**	**1.59***** **(5.93)**	**0.43***** **(3.97)**	**0.30***** **(2.79)**	**1.64***** **(6.11)**	**0.57***** **(5.90)**	**1.66***** **(6.33)**	**0.51***** **(4.47)**	**0.40***** **(3.51)**
market beta	-	-	0.19*** (10.59)	0.07*** (10.18)	0.07*** (11.88)	-	-	0.12*** (8.12)	0.05*** (8.59)	0.05*** (9.72)
ln(mv)	-	-	-0.18*** (-14.48)	-0.07*** (-13.28)	-0.06*** (-12.12)	-	-	-0.19*** (-16.67)	-0.08*** (-14.71)	-0.07*** (-13.72)
ln(mtb)	-	-	-0.00 (-0.09)	-0.01 (-1.44)	-0.04*** (-4.96)	-	-	0.02 (0.96)	-0.00 (-0.31)	-0.03*** (-3.41)
ln(turnov)	-	-	0.41*** (18.26)	0.10*** (10.93)	0.08*** (8.95)	-	-	0.36*** (16.94)	0.09*** (10.18)	0.08*** (8.38)
ln(spread)	-	-	0.28*** (15.90)	0.07*** (8.82)	0.07*** (8.99)	-	-	0.27*** (16.97)	0.08*** (10.02)	0.08*** (10.23)
lag-volatility	-	0.71*** (56.60)	-	0.62*** (41.93)	0.60*** (42.17)	-	0.69*** (50.12)	-	0.58*** (36.37)	0.56*** (37.57)
% Style Ownership	-	-	-	-	YES	-	-	-	-	YES
Other Controls	-	-	-	-	YES	-	-	-	-	YES
Adj-R^2 (%)	30.3	64.6	49.8	67.4	68.6	21.4	57.9	43.0	60.5	61.7
Number of observations	70,703	70,481	62,257	62,238	61,187	70,703	70,481	62,257	62,238	61,187

6. Econometric analysis and results: stock liquidity and style concentration.

6.1 Equation specification and control variables

We proceed with the analysis of the relation between the style concentration of ownership and three different measures of stock liquidity (price impact, bid-ask spread, share turnover). Since we want to test the predictive ability of style concentration on liquidity, we regress each of the liquidity variables of a stock at quarter q+1 on the H of quarter q:

$$\lambda_{i,q+1} = \alpha + \xi \cdot (H)_{i,q} + e_{i,q+1} \quad (6)$$

We expect that the style concentration is negatively correlated with the liquidity of the stock, thus our hypothesis is that:

$$H_0\!: \xi > 0 \;\; for\; ILLIQ\; \&\; bid - ask\; spread \quad OR \quad H_0\!: \xi < 0\; for\; turnover \quad (7)$$

We also include a number of control variables (Z) which are either known determinants of stock liquidity, or they are associated with the investment styles, or both. We also include the own lag value of stock liquidity, to address the endogeneity concerns, especially since liquidity measures are highly persistent:

$$\lambda_{i,q+1} = \alpha + \xi \cdot (H)_{i,q} + \Theta' \cdot Z_{i,q} + e_{i,q+1} \quad (8)$$

We include the natural logarithm of size as control variable to the regressions of all the liquidity variables. Size is a well-known proxy of liquidity, with the big stocks being much more liquid. The reason is that a bigger stock is followed and owned by a large number of different investors, it is thus much easier to find a counterpart. In addition, for a given size of trade, the impact to the price of a large stock will be much lower compared to the impact on the price of a small stock. Size is also a variable closely related with the definition of a large number of investment styles, thus we include it to isolate any effects attributable to any specific style. We also include the size in the specifications of share turnover (columns (4) and (5) of Table 8), as a style related control, since it is known that size and trading activity are not related.

We also include as control variable the idiosyncratic volatility of the stocks, since it is known that volatility is also closely related with liquidity. The price impact of volatile stocks is higher for a given quantity of traded shares. Furthermore we include share turnover as a control variable, because trading activity is related with the cost of trade. Investors could find faster and easier a counterpart for a stock with higher trading activity. In the column (5) of

Table 8 we also include turnover as an autoregressive variable, to address possible endogeneity problems, especially since turnover is a persistent variable.

Endogeneity concerns are present also for the cases of ILLIQ and bid-ask spread, which exhibit high serial correlation. It is possible that highly illiquid stocks are preferred only by funds that have expertise on their style. However, the expertise to certain stocks it is not closely related with the style, as style is much more general. Nevertheless, we include the lag of the respective values of ILLIQ and bid-ask spread, to address to a certain point the endogeneity concerns.

The natural logarithm of market-to-book value and the price momentum are also included as control variables to some of our specifications, to account for the value and the momentum dimensions of the style definitions. Momentum is also related to liquidity through the increased interest that creates to some stocks, which in turn increases their liquidity.

Moreover, we include to some of our specifications a host of risk related variables to control for any effect of risk on the liquidity of the stocks. Besides the four betas of the Carhart model (the three Fama-French factors plus a momentum one), we include the illiquidity beta. Amihud (2002) and Acharya and Pedersen (2005) find that stocks with high levels of illiquidity risk tend to be illiquid themselves. We thus include the illiquidity beta to control for this effect.

Finally, we include the shares of ownership of the individual styles, to further control for any style related effect. It is possible that stocks with increased levels of ownership from certain styles to be more illiquid or to have higher levels of trading activity. The inclusion of the shares of individual styles into the specification ensures that the coefficient of H shows the effect of style concentration and it is not affected by any specific style.

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.[65] The stock liquidity is serially correlated from quarter to quarter, thus we address the presence of firm effect (Cochrane (2001), Petersen (2009)) by estimating t-statistics that are based on robust (heteroskedasticity-consistent) standard errors, clustered in the firm level. This correction addresses both the heteroskedasticity and the serial correlation, which are present in our data, and corrects

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

(reduces) the size of the reported t-statistics. For easiness of exposition, we use three asterisks (***) to denote statistical significance at the 1% level, two asterisks (**) at the 5% level, and one asterisk (*) at the 10% level.

Due to the fact that liquidity measures (both ILLIQ and bid-ask spread), on average, decreased to the half of their magnitude after the first years of our sample, for each year we normalize the liquidity according to its cross-sectional mean and standard deviation. This way, the coefficient of the pooled-OLS regressions are meaningful and show the relation between liquidity and explanatory variables, clear from any systematic shifts to liquidity due to exogenous reasons. In addition, the coefficients of the regressions of ILLIQ and bid-ask spread, are directly expressed into standard deviations of the respective dependent variable. The series of share turnover do not exhibit any systematic shift to their values, thus we use the raw values.

The results (Tables 7 and 8) confirm our hypothesis, that higher style concentration is associated with lower stock liquidity. In all the specifications that we test, H enters into the regressions with a positive (negative for share turnover) and statistically significant coefficient (at 1%). The results are also economically significant, with one standard deviation of H predicting almost half standard deviation higher ILLIQ, almost one third standard deviation higher bid-ask spread and almost 0.13% lower daily share turnover (the average share turnover is 0.79%), in the univariate cases.

Table 7 reports the results of the regressions of a series of specifications regarding the relation between the ILLIQ variable and the style concentration. The univariate case is shown in column (1). A very strong relation is shown in this case. However, H is related with a number of known determinants of ILLIQ, thus in order to see which part of the relation is due to style concentration itself, we test the specification of column (2), in which we include the ln(size), the idiosyncratic volatility, the ln(turnover) and the ln(mtb). The coefficient of H remains statistically significant, although with a tenfold reduction of its magnitude. Nevertheless, again the effect of H is economically significant, with one standard deviation of H predicting 5% of a standard deviation higher ILLIQ.

In column (3) of Table 7 we include the lag of ILLIQ to the univariate specification. Liquidity is persistent and this is a way to address the possible endogeneity. The coefficient of H remains statistically significant, although further reduced. The fact that the coefficient of H remains significant means that the style concentration is a predictor of liquidity over and above any endogeneity concerns, since it has predicting power over the innovations of ILLIQ, which is rather difficult to affect the ex ante value of H. The coefficient is very small compared to that of column (1), but the interpretation now should be given relative to the size of the innovations of ILLIQ. One standard deviation of the innovations of ILLIQ is 0.17 standard deviations of ILLIQ level. Thus one standard deviation of H predicts roughly 5% of the standard deviation of the innovations of ILLIQ.

The column (4) of Table 7 is the combination of columns (2) and (3). We simultaneously control for the lag value of ILLIQ and for its known determinants. The coefficient of H remains statistically significant (t-statistic 8.88) and is even higher than that of column (3), indicating that one standard deviation of H explains roughly 12% of the standard deviation of the innovations of ILLIQ. In column (10) we use the full specification case including the full set of control variables and the ownership share of each style. Again the coefficient of H remains highly significant with t-statistic equals 6.89. The effect of style concentration is robust to the inclusion of any related control variable, even the lag value of ILLIQ itself.

The coefficient of the control variables is as expected. Size and share turnover enter into the regressions with negative coefficients and volatility with a positive one. The coefficient of market-to-book change sign from positive to negative when we include the lag value of ILLIQ, showing that given the ex ante level of liquidity, a "glamour" stock is more liquid. The autoregressive coefficient of ILLIQ is 0.98 when only H is the additional variable, and it falls around 0.70 in the full specification case. In unreported results we find that the inclusion of additional lags of ILLIQ does not change a lot to the results. In Appendix we also show the results for the rest of the control variables.

Table 8 reports the results of the regressions of a series of specifications regarding the relation of the bid-ask spread (columns (1) to (3)) and the share turnover (columns (4) and (5)) with the style concentration. Column (1) shows the results of the univariate regression of the bid-ask spread on the H of the previous quarter. The result is strong both in statistical and economic sense.

Bid-ask spread also exhibits high serial correlation and endogeneity concerns – similar to those related to ILLIQ- are present. In order to address these concerns we include to our specification the lag value of bid-ask spread (column (2) of Table 8). The coefficient of H remains statistically significant (t-statistic 9.85), however there is a tenfold reduction in its magnitude. Nevertheless, this is the effect on the change of the bid-ask spread, which has a standard deviation of 0.43 standard deviations of the level of spread. Thus, one standard deviation of H predicts almost 7% of the standard deviation of the innovations of the spread, which is economically significant.

The column (3) of Table 8 reports the results of the regression of bid-ask spread on H and the full host of control variables. The coefficient of H remains statistically significant (t-statistic 4.14) and its coefficient is slightly reduced from that of column (2). The coefficients of the control variables are the expected ones (the coefficients of the rest of control variables are reported on Appendix) and the autoregressive term of bid-ask spread is around 0.85.

The column (4) shows the result of the univariate regression of share turnover on H. The result is very strong indicating that stocks with ownership that is populated by owners of only few styles, exhibit much less trading activity. This is the expected result, since stocks with pluralistic ownership in terms of style orientation change hands more frequently, due to style investing described by Barberis and Shleifer (2003). Column (5) shows the results of the regression with the complete set of controls. Again, the interpretation of the coefficient is on terms of innovations of turnover, due to the inclusion of lag share turnover (autoregressive coefficient = 0.74). The coefficient of H remains statistically significant (t-statistic -3.95) and one standard deviation of H predicts around 0.05% lower change on average daily turnover, which is economically significant (the average change in daily turnover is orders of magnitude smaller). The coefficient of bid-ask spread is negative and significant, indicating that more liquid stocks exhibit higher trading activity. Size, book-to-market and volatility enter the regressions with insignificant coefficients. The results for the rest of control variables are reported in Appendix, with the coefficients of illiquidity risk, momentum and momentum-beta to be significant and to the right direction.

Our results are in line with that of Amihud et al. (1999) and of Rubin (2007), who also find that the concentration of ownership increases the liquidity of a stock. They are also in line with the predictions of Allen and Gale (1994) and of Duffie and Strulovici (2012), regarding the higher price impacts due to reduced participation to the market (or to a specific stock). Mitchell et al. (2007) also find evidence that reduced presence of capital to a specific market may result in large price impacts, due to the slow moving of capital from market to market. Our finding of negative relation between style concentration of ownership and stock liquidity is closely connected to our finding on the positive relation of concentration with stock price volatility.

Our paper contributes to the style investing literature as it is the first to show that style investing affects the liquidity of a stock, through the segmentation of ownership according to styles. A stock with a lot of different styles present to its ownership is more likely to exhibit higher price impacts, because at any specific moment, investors with different preferences are present and ready to accommodate any excess demand (or to absorb any excess supply) of the stock.

Table 7: Stock price impact (ILLIQ) and previous quarter's style concentration in ownership

Panel OLS regressions of the ln(ILLIQ) of stock i at quarter q+1, $\lambda_{i,q+1}$, 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:

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

There are 5 regressions in columns 1 through 5. 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, betas of SMB, HML and MOM factors, illiquidity beta, price momentum and ln(dta)). See Table 1 for the detailed definitions of the variables.
The natural logarithm of ILLIQ is normalized in each quarter; the coefficients are expressed in terms of standard deviations of ln(ILLIQ). 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 robust standard errors that are clustered by firm. 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)
H	**4.17*** (20.28)**	**0.42*** (10.18)**	**0.06*** (5.81)**	**0.17*** (8.88)**	**0.13*** (6.89)**
ln(mv)	-	-0.46*** (-223.42)	-	-0.12*** (-22.68)	-0.13*** (-25.22)
idiosyncratic volatility	-	9.43*** (24.10)	-	1.68*** (9.58)	1.89*** (10.65)
ln(turnover)	-	-0.40*** (-54.19)	-	-0.09*** (-22.70)	-0.04*** (-4.96)
ln(mtb)	-	0.01** (2.36)	-	-0.01*** (-6.34)	-0.01*** (-5.50)
lag-ILLIQ	-	-	0.98*** (1046.55)	0.72*** (62.21)	0.70*** (62.45)
% Style Ownership	-	-	-	-	YES
Other Controls	-	-	-	-	YES
Adj-R^2 (%)	19.5	96.2	97.1	97.5	97.7
Number of observations	70,714	67,105	70,509	67,102	63,751

Table 8: Bid-ask spread, share turnover and previous quarter's style concentration in ownership

Panel OLS regressions of the average natural logarithm of daily bid-ask spread (or of the average daily share turnover) of stock i at quarter q+1, $\lambda_{i,q+1}$, 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:

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

There are 5 regressions in columns 1 through 5. Columns (1) to (3) show the results for the bid-ask spread and columns (4) and (5) show the results for the share turnover. 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, betas of SMB, HML and MOM factors, illiquidity beta, price momentum and ln(dta)). See Table 1 for the detailed definitions of the variables.
Bid-ask spread is expressed in percentage over stock price and its natural logarithm is normalized for each quarter; the coefficients are expressed in terms of standard deviations of ln(spread) (Columns (1) to (3)). Share turnover is expressed in percentage terms. The sample covers the period between 1997-Q1 and 2015-Q4 (76 quarters) and on average consists of around 772 stocks in each quarter for bid-ask spread and 805 stocks in each quarter for share turnover. 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 robust standard errors that are clustered by firm. 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)
	bid-ask spread	bid-ask spread	bid-ask spread	turnover	turnover
H	**2.56*** (15.98)**	**0.26*** (9.85)**	**0.15*** (3.82)**	**-1.09*** (-17.43)**	**-0.44*** (-4.06)**
ln(mv)	-	-	-0.03*** (-15.08)	-	-0.01 (-0.80)
ln(mtb)	-	-	-0.02*** (-8.59)	-	-0.11 (-1.22)
idiosyncratic volatility	-	-	1.87*** (6.40)	-	1.04 (0.44)
lag-turnover	-	-	-0.26** (-2.16)	-	0.74*** (18.09)
ln(spread)	-	0.88*** (270.39)	0.83*** (144.87)	-	-0.04*** (-3.29)
% Style Ownership	-	-	YES	-	YES
Other Controls	-	-	YES	-	YES
Adj-R^2 (%)	7.5	80.9	80.6	1.8	57.9
Number of observations	67,868	64,919	58,704	69,824	61,186

7. Robustness of the econometric results

7.1 The effect of the Financial Crisis of 2007-2009

Our sample includes the quarters of the financial crisis of 2007-2009, during which the relations between the variables may change considerably. It is possible that stocks with high style concentration of ownership become very illiquid and volatile during the quarters of the crisis, affecting our results. In order to check the robustness of our results we run again the regressions of sections 5 and 6, this time excluding the 8 quarters of the crisis (3Q2007-2Q2009).

Table 9 reports the results of the regressions of total and idiosyncratic volatility on H and other control variables (it is the respective of Table 6). The results are almost identical, with the coefficient of H becoming slightly bigger and statistically stronger. The coefficients of market beta, ln(mtb), ln(spread) and the autoregressive term becomes slightly smaller and weaker (the autoregressive term lose around 20% of its statistical power). It seems that during the crisis the increased persistence of volatility (volatility clustering) and the increased role of liquidity, explain more of the variability of volatility, in expense of the predictive power of H. In addition, these results indicate that higher style concentration does not predict any increase in stock volatility, during extreme events. Characteristics related to the risk and the liquidity of stocks seem to be more relevant to explain the within the crisis changes of volatility.

Tables 10 and 11 report the results of the regressions of ILLIQ, bid-ask spread and share turnover on H and the rest of control variables (they are the respective of Tables 7 and 8). Again, the results are very close to the results of the complete sample. The coefficient of H is slightly weaker compared to the full period regressions. It seems that as for the volatility, the predictive ability is not driven from the crisis quarters. However, the (slight) weakening of the coefficient of H (for all the three liquidity variables) indicates that during the crisis stocks with higher H tend to be even more illiquid. Nevertheless, this effect is very small and does not drive the whole results.

Table 9: Stock price volatility and previous quarter's style concentration in ownership – Exclusion of the quarters of the Financial Crisis (3Q2007- 2Q2009).

Panel OLS regressions of the daily price volatility of stock i at quarter q+1, $\sigma_{i,q+1}$, 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:

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

There are 10 regressions in columns 1 through 10. Columns (1) to (5) show the results for the total volatility and columns (6) to (10) show the results for the idiosyncratic volatility (the residual of a 4-factor model (Fama-French, Carhart)). A time effect with quarterly dummies is included in every regression. The variables of each regression are described in the very left column. Lag-volatility is the lagged total volatility for columns (1) to (5) and the lagged idiosyncratic volatility for columns (6) to (10). 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: betas of SMB, HML and MOM factors, illiquidity beta, price momentum and ln(dta)). See Table 1 for the detailed definitions of the variables.

Volatility is 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 814 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 robust standard errors that are clustered by firm. 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)
	total	total	total	total	total	idio/tic	idio/tic	idio/tic	idio/tic	idio/tic
H	**1.19*** (4.51)**	**0.41*** (4.44)**	**1.63*** (6.50)**	**0.46*** (4.31)**	**0.33*** (3.10)**	**1.62*** (6.01)**	**0.55*** (5.68)**	**1.75*** (6.97)**	**0.55*** (4.87)**	**0.44*** (3.79)**
market beta	-	-	0.15*** (9.90)	0.06*** (9.42)	0.07*** (12.37)	-	-	0.10*** (7.44)	0.05*** (8.14)	0.05*** (10.28)
ln(mv)	-	-	-0.18*** (-15.56)	-0.07*** (-13.09)	-0.06*** (-12.39)	-	-	-0.20*** (-17.36)	-0.08*** (-14.08)	-0.07*** (-13.71)
ln(mtb)	-	-	0.05*** (2.65)	0.01 (1.16)	-0.02*** (-2.91)	-	-	0.05*** (2.90)	0.01 (1.36)	-0.02** (-2.27)
ln(turnov)	-	-	0.39*** (17.73)	0.10*** (10.47)	0.08*** (8.79)	-	-	0.35*** (16.62)	0.09*** (9.81)	0.08*** (8.03)
ln(spread)	-	-	0.23*** (13.66)	0.06*** (7.72)	0.06*** (7.85)	-	-	0.22*** (14.80)	0.07*** (8.51)	0.06*** (8.58)
lag-volatility	-	0.70*** (45.18)	-	0.61*** (33.11)	0.58*** (35.35)	-	0.69*** (41.32)	-	0.58*** (29.71)	0.55*** (31.43)
% Style Ownership	-	-	-	-	YES	-	-	-	-	YES
Other Controls	-	-	-	-	YES	-	-	-	-	YES
Adj-R^2 (%)	17.4	57.3	40.2	60.1	61.7	16.5	55.0	39.4	57.7	59.2
Number of observations	63,484	63,305	56,400	56,389	55,398	63,484	63,305	56,400	56,389	55,398

Table 10: Stock price impact (ILLIQ) and previous quarter's style concentration in ownership – Exclusion of the quarters of the Financial Crisis (3Q2007- 2Q2009).

Panel OLS regressions of the ln(ILLIQ) of stock i at quarter q+1, $\lambda_{i,q+1}$, 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:

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

There are 5 regressions in columns 1 through 5. 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, betas of SMB, HML and MOM factors, illiquidity beta, price momentum and ln(dta)). See Table 1 for the detailed definitions of the variables.

The natural logarithm of ILLIQ is normalized in each quarter; the coefficients are expressed in terms of standard deviations of ln(ILLIQ). 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 841 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 robust standard errors that are clustered by firm. 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)
H	**4.13*** (20.58)**	**0.42*** (10.18)**	**0.06*** (5.67)**	**0.16*** (8.28)**	**0.12*** (6.25)**
ln(mv)	-	-0.45*** (-218.84)	-	-0.12*** (-21.75)	-0.13*** (-24.97)
idiosyncratic volatility	-	10.91*** (23.80)	-	1.97*** (9.94)	2.34*** (12.29)
ln(turnover)	-	-0.40*** (-56.22)	-	-0.09*** (-21.67)	-0.10*** (-24.25)
ln(mtb)	-	0.01** (2.31)	-	-0.01*** (-6.32)	-0.01*** (-5.24)
lag-ILLIQ	-	-	0.98*** (997.70)	0.72*** (60.42)	0.71*** (62.48)
% Style Ownership	-	-	-	-	YES
Other Controls	-	-	-	-	YES
Adj-R^2 (%)	19.7	96.2	97.2	97.5	97.7
Number of observations	63,500	60,148	63,322	60,145	57,202

Table 11: Bid-ask spread, share turnover and previous quarter's style concentration in ownership – Exclusion of the quarters of the Financial Crisis (3Q2007- 2Q2009).

Panel OLS regressions of the average natural logarithm of daily bid-ask spread (or of the average daily share turnover) of stock i at quarter q+1, $\lambda_{i,q+1}$, 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:

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

There are 5 regressions in columns 1 through 5. Columns (1) to (3) show the results for the bid-ask spread and columns (4) and (5) show the results for the share turnover. 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, betas of SMB, HML and MOM factors, illiquidity beta, price momentum and ln(dta)). See Table 1 for the detailed definitions of the variables.
Bid-ask spread is expressed in percentage over stock price and its natural logarithm is normalized for each quarter; the coefficients are expressed in terms of standard deviations of ln(spread) (Columns (1) to (3)). Share turnover is expressed in percentage terms. 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 on average consists of around 778 stocks in each quarter for bid-ask spread and 814 stocks in each quarter for share turnover. 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 robust standard errors that are clustered by firm. 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)
	bid-ask spread	bid-ask spread	bid-ask spread	turnover	turnover
H	**2.30*** (14.98)**	**0.25*** (9.57)**	**0.13*** (3.27)**	**-0.99*** (-15.79)**	**-0.47*** (-3.71)**
ln(mv)	-	-	-0.03*** (-13.97)	-	-0.01 (-0.94)
ln(mtb)	-	-	-0.02*** (-7.98)	-	-0.11 (-1.14)
idiosyncratic volatility	-	-	2.31*** (6.49)	-	0.65 (0.24)
lag-turnover	-	-	-0.19*** (-2.62)	-	0.74*** (17.56)
ln(spread)	-	0.88*** (256.65)	0.83*** (142.57)	-	-0.03*** (-2.86)
% Style Ownership	-	-	YES	-	YES
Other Controls	-	-	YES	-	YES
Adj-R^2 (%)	6.3	79.7	79.6	1.3	57.4
Number of observations	60,655	58,580	52,915	62,608	55,397

7.2 The effect of outliers

We furthermore check for the effect of the outliers to our results. We thus winsorize all the variables (except from the shares of the 32 styles) to the 0.5% of each tail of their distribution and re-run all the regressions. Tables 12, 13 and 14 report the results of the regressions with the data after the winsorization.

Table 12 shows the results of the regressions of total and idiosyncratic volatility on H and the rest of the control variables (it is the respective of Table 6). There is not any considerable change on the results. The coefficient of H is slightly smaller but at similar levels of statistical significance. The most significant change is on the statistical significance of the autoregressive term of volatility (both of the total and of the idiosyncratic), which has doubled t-statistics after the winsorization. It seems that cases of extreme volatility do not depend on the level of the stock volatility, but they rather be attributable to special events.

Table 13 shows the results of the regressions of ILLIQ on H and the rest control variables (it is the respective of Table 7). The results of Table 13 indicate that outliers do not drive the results of the regressions of ILLIQ. The coefficient of H remains at the same level and with t-statistics very close to that of Table 7.

Finally, Table 14 reports the results of the regressions of bid-ask spread and of share turnover on H and other control variables (it is the respective of Table 8). The coefficient of H remain positive and significant (t-statistic 2.96) in the full specification case for bid-ask spread (column (3)), although its value reduced to 0.11 from 0.16 of Table 8. The coefficient of H in the full specification case for the share turnover (column (5)) is smaller in size (-0.14 from -0.44 on Table 8), but stronger in significance (t-statistic -6.15 from -3.95 on Table 8).

Table 12: Stock price volatility and previous quarter's style concentration in ownership – Winsorization of the variables at the 0.5% of each tail of their distribution.

Panel OLS regressions of the daily price volatility of stock i at quarter q+1, $\sigma_{i,q+1}$, 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:

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

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. Columns (1) to (5) show the results for the total volatility and columns (6) to (10) show the results for the idiosyncratic volatility (the residual of a 4-factor model (Fama-French, Carhart)). A time effect with quarterly dummies is included in every regression. The variables of each regression are described in the very left column. Lag-volatility is the lagged total volatility for columns (1) to (5) and the lagged idiosyncratic volatility for columns (6) to (10). 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: betas of SMB, HML and MOM factors, illiquidity beta, price momentum and ln(dta)). See Table 1 for the detailed definitions of the variables.
Volatility is measured in percentage form. The sample covers the period between 1997-Q1 and 2015-Q4 (76 quarters) and on average consists of around 805 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 robust standard errors that are clustered by firm. 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)
	total	total	total	total	total	idio/tic	idio/tic	idio/tic	idio/tic	idio/tic
H	**1.03*** (4.37)**	**0.33*** (4.74)**	**1.52*** (6.63)**	**0.33*** (4.22)**	**0.24*** (2.70)**	**1.46*** (6.17)**	**0.45*** (6.25)**	**1.56*** (7.09)**	**0.38*** (4.71)**	**0.32*** (3.42)**
market beta	-	-	0.20*** (16.25)	0.06*** (13.18)	0.07*** (13.47)	-	-	0.14*** (12.01)	0.05*** (10.56)	0.05*** (10.28)
ln(mv)	-	-	-0.18*** (-15.35)	-0.06*** (-15.09)	-0.06*** (-13.12)	-	-	-0.19*** (-17.66)	-0.07*** (-17.09)	-0.06*** (-14.92)
ln(mtb)	-	-	0.01 (0.81)	-0.01 (-0.69)	-0.03*** (-4.61)	-	-	0.03** (1.97)	0.00 (0.65)	-0.02*** (-2.71)
ln(turnov)	-	-	0.42*** (20.16)	0.08*** (13.38)	0.07*** (10.43)	-	-	0.36*** (18.71)	0.08*** (11.87)	0.06*** (9.38)
ln(spread)	-	-	0.27*** (17.09)	0.06*** (10.22)	0.06*** (10.20)	-	-	0.27*** (18.46)	0.07*** (11.86)	0.07*** (11.77)
lag-volatility	-	0.74*** (119.20)	-	0.65*** (83.11)	0.64*** (83.10)	-	0.72*** (110.54)	-	0.62*** (75.03)	0.60*** (73.51)
% Style Ownership	-	-	-	-	YES	-	-	-	-	YES
Other Controls	-	-	-	-	YES	-	-	-	-	YES
Adj-R^2 (%)	32.8	69.7	53.5	67.4	72.5	24.0	64.1	47.1	65.7	66.6
Number of observations	70,703	70,481	62,257	62,238	61,187	70,703	70,481	62,257	62,238	61,187

Table 13: Stock price impact (ILLIQ) and previous quarter's style concentration in ownership – Winsorization of the variables at the 0.5% of each tail of their distribution.

Panel OLS regressions of the ln(ILLIQ) of stock i at quarter q+1, $\lambda_{i,q+1}$, 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:

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

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 5 regressions in columns 1 through 5. 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, betas of SMB, HML and MOM factors, illiquidity beta, price momentum and ln(dta)). See Table 1 for the detailed definitions of the variables.

The natural logarithm of ILLIQ is normalized in each quarter; the coefficients are expressed in terms of standard deviations of ln(ILLIQ). 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 robust standard errors that are clustered by firm. 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)
H	**4.12*** (20.43)**	**0.43*** (8.31)**	**0.06*** (5.63)**	**0.17*** (8.32)**	**0.13*** (6.64)**
ln(mv)	-	-0.45*** (-175.08)	-	-0.12*** (-15.43)	-0.13*** (-25.46)
idiosyncratic volatility	-	11.15*** (33.68)	-	1.96*** (10.69)	2.48*** (17.58)
ln(turnover)	-	-0.41*** (-42.94)	-	-0.09*** (-19.15)	-0.10*** (-28.04)
ln(mtb)	-	0.01** (2.21)	-	-0.01*** (-6.76)	-0.01*** (-5.59)
lag-ILLIQ	-	-	0.98*** (1077.59)	0.73*** (45.90)	0.69*** (61.26)
% Style Ownership	-	-	-	-	YES
Other Controls	-	-	-	-	YES
Adj-R^2 (%)	19.3	96.1	97.2	97.5	97.7
Number of observations	70,714	67,105	70,509	67,102	63,751

Table 14: Bid-ask spread, share turnover and previous quarter's style concentration in ownership – Winsorization of the variables at the 0.5% of each tail of their distribution.

Panel OLS regressions of the average natural logarithm of daily bid-ask spread (or of the average daily share turnover) of stock i at quarter q+1, $\lambda_{i,q+1}$, 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:

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

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 5 regressions in columns 1 through 5. Columns (1) to (3) show the results for the bid-ask spread and columns (4) and (5) show the results for the share turnover. 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, betas of SMB, HML and MOM factors, illiquidity beta, price momentum and ln(dta)). See Table 1 for the detailed definitions of the variables.

Bid-ask spread is expressed in percentage over stock price and its natural logarithm is normalized for each quarter; the coefficients are expressed in terms of standard deviations of ln(spread) (Columns (1) to (3)). Share turnover is expressed in percentage terms. The sample covers the period between 1997-Q1 and 2015-Q4 (76 quarters) and on average consists of around 772 stocks in each quarter for bid-ask spread and 805 stocks in each quarter for share turnover. 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 robust standard errors that are clustered by firm. 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)
	bid-ask spread	bid-ask spread	bid-ask spread	turnover	turnover
H	**2.47*** (16.17)**	**0.25*** (10.16)**	**0.10*** (2.71)**	**-1.08*** (-19.64)**	**-0.13*** (-5.38)**
ln(mv)	-	-	-0.03*** (-15.10)	-	-0.01*** (-2.62)
ln(mtb)	-	-	-0.02*** (-8.65)	-	-0.00 (-0.15)
idiosyncratic volatility	-	-	3.05*** (9.12)	-	-1.59*** (-5.91)
lag-turnover	-	-	-3.02*** (-6.69)	-	0.87*** (120.58)
ln(spread)	-	0.88*** (279.03)	0.83*** (143.88)	-	-0.01*** (-3.37)
% Style Ownership	-	-	YES	-	YES
Other Controls	-	-	YES	-	YES
Adj-R^2 (%)	7.4	80.6	80.5	16.0	80.3
Number of observations	67,868	64,919	58,704	69,824	61,186

8. Conclusion

A number of new studies shows that the ownership structure of stocks determines to a large extend the volatility and the liquidity of stocks, due to the correlation of the trading needs of the stock owners. In addition, the amount of capital that is readily available for each stock determines the extend of the price impact of a stock due to excess demand/supply, or put it differently, how flat is the demand curve for each stock, or how much is a stock exposed to the limits-of-arbitrage.

The self-definition of institutional investors into different styles (and types) characterizes the institutional ownership (which is now very high, close to 80% on average for NYSE stocks). The stylization of the stock market creates both the abovementioned phenomena: limited availability of readily available capital for stocks and correlated trading needs of the stock owners. We thus should be able to detect a relationship between the exposure of a stock to the stylization of the stock market and its volatility and liquidity.

We follow an innovative way to measure the exposure of a stock to the stylization of the market. Instead of classify each stock to one style, we have a more realistic approach and measure the concentration of different styles to the ownership of each stock. We thus avoid assuming that each stock belongs to one style, since the data show that each stock is held by a number of different styles. Style concentration is a measure of the exposure of each stock to the stylization of the market. We subsequently formulate the hypothesis that higher style concentration is related with higher stock price volatility and lower liquidity.

Our results confirm our hypothesis and show that there is a strong relation between the style concentration in ownership and the volatility and liquidity of the stocks. We show that on the univariate cases, one standard deviation of H predicts 13.5 b.p. (19.3 b.p.) higher future daily total (idiosyncratic) volatility, 0.5 (0.3) standard deviations higher ILLIQ (bid-ask spread) and 0.13% lower daily share turnover. The effect of style concentration is robust for any specification we use (and for the inclusion of the lag of the respective dependent variable). We also show that our results are not driven by the financial crisis of 2007-2009 nor from the outliers.

Our results add to the findings of the literature that study the relation between ownership structure and volatility (Greenwood and Thesmar, 2011) and liquidity (Amihud et al (1999), Rubin (2007)). Our result is complementary to theirs, as we show that the ownership concentration in a higher level of aggregation does also affect volatility and liquidity. We also contribute to the literature which studies the effect of the institutional ownership to stock risk and volatility (Benett et al., 2003), as we show that not only the institutional ownership per se, but its structure as well is a determinant of volatility and liquidity. In addition, we contribute to the literature of style investing (Barberis and Shleifer (2003), Teo and Woo (2004), Froot and Teo (2008), Wahal and Yavuz (2013)), introducing a new way to measure the exposure of stock to the stylization and relaxing the assumption that each stock belongs to only one style. Our approach enables us to study the effect of style investing in volatility and liquidity, and show that style investing not only creates momentum, reversal and comovement, but price volatility and stock liquidity as well.

Future research could study the relation between style concentration and the comovement of stocks, either due to the market comovement, or due to risk-factor comovement, or the style oriented comovement. In addition, the relation between the style concentration and the exposure of stocks to a series of risk (market risk, volatility of the market returns, market illiquidity) could be studied. Finally, one could extend the analysis to the marketwide level, measuring the relation between marketwide changes in style concentration and market volatility – market liquidity.

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:

33) **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.

34) **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.

35) **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.

36) **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.

37) **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.

38) **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.

39) **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.

40) **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.

41) **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.

42) **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.

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

44) **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.

45) **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.

46) **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.

47) **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.

48) **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.

49) **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.

50) **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.

51) **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.

52) **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.

53) **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.

54) **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.

55) **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.

56) **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.

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

58) **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.

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

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

61) **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.

62) **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.

63) **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.

64) **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: Regression coefficients of the rest control variables

Table 15: Results of rest control variables, from Tables 6 (columns (5) and (10)), 7 (column (5)) and 8 (columns (3) and (5)).

Table – Column	Table 6 Column (5)	Table 6 Column (10)	Table 7 Column (5)	Table 8 Column (3)	Table 8 Column (5)
Dependent Variable	tot_vol	idio_vol	ILLIQ	spread	turnover
market beta	0.07***	0.05***	0.01***	-0.00	0.01
	(11.88)	(9.72)	(4.99)	(-0.94)	(0.92)
SMB beta	0.02***	0.01***	0.00***	0.00	-0.00
	(3.30)	(2.68)	(3.89)	(0.66)	(-0.30)
HML beta	0.00	-0.00	-0.00	0.00*	-0.01
	(0.26)	(-0.11)	(-0.81)	(1.86)	(-1.00)
MOM beta	-0.04***	-0.03***	-0.00	0.00	-0.01***
	(-7.67)	(-6.07)	(-0.74)	(0.64)	(-2.71)
illiq-beta	-0.00	0.00	-0.00	-0.00	-0.01**
	(-0.34)	(0.17)	(-1.39)	(-0.05)	(-2.11)
momentum	0.05***	0.01	-0.01***	-0.02***	0.09**
	(3.91)	(1.47)	(-2.72)	(-4.01)	(2.04)
ln(d-t-a)	-0.00**	-0.00**	-0.00***	0.00	-0.00
	(-2.05)	(-2.05)	(-2.67)	(0.56)	(-0.47)

References

Acharya Viral V., P.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.

Allen, F., Gale, D., 1994. Limited Market Participation and Volatility of Asset Prices. The American Economic Review 84, 933-955

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

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

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

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

Boyer, B. H. (2011), Style-Related Comovement: Fundamentals or Labels? The Journal of Finance 66, 307–332.

Bushee, B.J., Noe, C.F., 2000. Corporate Disclosure Practices, Institutional Investors, and Stock Return Volatility. Journal of Accounting Research 38, 171-202

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

Coval, J., Stafford, E., 2007. Asset fire sales (and purchases) in equity markets. Journal of Financial Economics 86, 479-512

Duffie, D., Strulovici, B., 2012. Capital Mobility and Asset Pricing. Econometrica 80, 2469-2509

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

Frazzini, A., Lamont, O.A., 2008. Dumb money: Mutual fund flows and the cross-section of stock returns. Journal of Financial Economics 88, 299-322

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

Gabaix, X., Gopikrishnan, P., Plerou, V., Stanley, H.E., 2006. Institutional Investors and Stock Market Volatility. The Quarterly Journal of Economics 121, 461-504

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

Greenwood, R., & Vayanos, D. (2010). Price Pressure in the Government Bond Market. *The American Economic Review, 100*(2), 585-590.

Harris, L. and Gurel, E. (1986), Price and Volume Effects Associated with Changes in the S&P 500 List: New Evidence for the Existence of Price Pressures. The Journal of Finance 41, 815–829.

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

Kumar, A. (2009) 'Dynamic Style Preferences of Individual Investors and Stock Returns', *Journal of Financial and Quantitative Analysis*, 44(3), pp. 607–640.

Lou, D., 2012. A Flow-Based Explanation for Return Predictability. Review of Financial Studies 25, 3457-3489

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

Mitchell, M., Pedersen, L.H., Pulvino, T., 2007. Slow Moving Capital. The American Economic Review 97, 215-220

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

Rubin, A., 2007. Ownership level, ownership concentration and liquidity. Journal of Financial Markets 10, 219-248

Shleifer, A., 1986. Do Demand Curves for Stocks Slope Down? The Journal of Finance 41, 579-590

Sias, R.W., 1996. Volatility and the Institutional Investor. Financial Analysts Journal 52, 13-20

Sias, R.W., 2004. Institutional Herding. Review of Financial Studies 17, 165-206

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

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

Chapter 3: Liquidity and Stock Returns during Large Market Declines

1. Introduction

In this paper we empirically test the predictive role of stock illiquidity on stock returns during a large market decline. Theoretical and empirical work until now predicts a positive relation between liquidity and stock returns during a crisis (or a negative relation when an illiquidity measure instead of a liquidity one is used). We find that illiquidity does not have predictive ability over returns during the financial crisis of 2007-2009. This is in line with other empirical works which use data from this turmoil period and find that liquid assets also suffer big losses and that hedge funds mainly sell liquid stocks to meet their funding needs.

The novelty of this paper is the augmentation of a, previously undocumented, signed turnover factor (individual for each stock) into a standard four factor model (three factors of Fama – French plus a momentum factor) of stock returns. This signed turnover factor (henceforth STF) is constructed in a weekly basis as the sum of the daily turnover (to which the sign of the corresponding daily return is given) for the trading days of each weekly period. STF captures the impact of non-fundamental demand (e.g. fund flows) for stocks, on the returns.

During the crisis period we split STF into a normal and an abnormal part. We assume that the abnormal part represents the excess need for trading (selling in our case) for reasons related with the shock to the market (funding liquidity problems, redemptions of funds, internal risk models, wealth effects, all these mechanisms are described on the limits of arbitrage literature). When we use our augmented factor model to measure the "normal" returns during the crisis, we use as input of the extra factor only the normal part of the STF. Thus, the remaining unexplained returns (the abnormal returns) include the effect of abnormal trading (selling) during the crisis.

Abnormal returns also include the effect of flight to liquidity, a substitution effect towards liquid stocks due to investors' preference to hold these stocks during a crisis (to avoid large negative price impacts). If we use the sum of normal and abnormal part of STF for estimating of the normal returns, we assume that the whole trading activity of the crisis period is normal (and expected) and ignore that investors may choose to sell stocks minimizing the price impact. An explanation could be that they prefer selling liquid stocks, in accordance to the stylized facts of the last financial crisis[66]. In this way, we induce a bias to the measurement of abnormal returns, in favor of liquid stocks.

In our main econometric analysis, we run a cross-sectional regression of the 26-week cumulative abnormal returns (CAR) (19 and 26 weeks after the collapse of Lehman Brothers) on the individual level of illiquidity (as it was measured 3 months before the event date). As a measure of illiquidity we use ILLIQ (Amihud, 2002) which is a widely used illiquidity measure, with high correlation with both high-frequency measures of illiquidity and other low frequency measures.[67] The coefficient of illiquidity in the regression is not significant, indicating that illiquidity has limited predicting ability on stock returns during a crisis period. When we use the bid-ask spread as the main independent variable, results remain unchanged.

We subsequently split the cumulative abnormal returns into two parts, the part induced by the abnormal STF (cumulative flight-from-liquidity part, CFFL) and the part induced by substitution effects towards liquid stocks (cumulative flight-to-liquidity part, CFTL). We separately run cross-sectional regressions of the two parts on the level of illiquidity. Our results reveal that illiquidity affects stock returns during a crisis with two distinct and opposite directions, which partially offset each other.

When we use as dependent variable the part that is related with the abnormal STF (CFFT), illiquidity has a positive and statistically significant coefficient. That is, illiquid stocks perform better than liquid stocks, during the period after the Lehman’s event. This is an interesting result which we call it flight-from-liquidity, as we interpret it as the effect of selling liquid stocks by the investors to absorb funding liquidity, during a crisis. On the contrary, when we use as dependent variable the part that is related with the substitution effects towards liquid stocks (CFTL), illiquidity has a negative and statistically significant coefficient. This means that illiquid stocks perform worse than the liquid ones, during the large market decline after the Lehman’s default. This result is in line with the flight-to-liquidity prediction.

We also conduct a whole period analysis, to confirm that the abovementioned relations are the result of the large market decline, and exclude a possible permanently present in the stock market. We, thus, estimate rolling weekly cross-sectional regressions of the 26-week ahead cumulative abnormal returns (CAR) and of their two parts (CFFL and CFTL) on

---

[66] Scholes (2000) starts the debate about the choice of an investor who needs to liquidate part of his portfolio and his portfolio includes both liquid and illiquid assets.

[67] Goyenko et al. (2009), Hasbrouck (2009). Besides, the use of a price impact measure is preferable over a spread measure, as the latter captures the cost of transacting for small quantities.

the ILLIQ of the stocks. Our results show that for the period 2004-2007, the main driver of the illiquidity premium is the CFTL part, while this changes during the financial crisis of 2007-2009. Flight-from-liquidity is a phenomenon that manifests selectively during the crisis.

We furthermore decompose ILLIQ to volatility and size, its two basic components, to specify the drivers of the two opposite effects (flight-to-liquidity and flight-from-liquidity). When we repeat our analysis using as main independent variables the size and the volatility instead of the ILLIQ we find that size is connected with the flight-from-liquidity effect, while volatility is connected with the flight-to-liquidity effect. Overall, the effect of size and the effect of volatility partially offset each other.

To summarize, our results indicate that individual stock liquidity is not a good predictor of the stock returns during a large market decline. Contrary to the predictions of the literature, about the flight-to-liquidity phenomenon, we show that there is a simultaneous co-existence of flight-to-liquidity and flight-from-liquidity phenomena, which affect the returns and partially offset each other. We also find that liquidity risk is a poor predictor of stock returns during the crisis. The reason for this failure of illiquidity risk to predict stock returns during the crisis is its large cross-sectional correlation with illiquidity per se.

The remainder of the paper is organized as follows. In section 2 we review the related literature and explain the contribution of this study. In section 3 we describe our empirical methodology. In section 4 we describe the data. In section 5 we describe the variables. In section 6 we validate that STF is a proxy for non-fundamental demand for stocks. In section 7 we illustrate the results of the whole period analysis. In section 8 we report the basic statistics of the independent variables of the main econometric analysis. In section 9 we illustrate the econometric results of the event period. In section 10 we provide the results of the ILLIQ decomposition. In section 11 we report the results of a robustness check and finally in section 12 we conclude.

2. Related literature

Our paper is part of the literature that examines the role of illiquidity during the financial crisis of 2007-2009. Ben-David et al. (2011) find that hedge funds absorbed liquidity to meet their funding needs (due to redemptions and forced deleveraging) by selling liquid stocks and high-volatility stocks (simultaneous existence of flight-from-liquidity and flight-to-liquidity). Also, Jotikasthira et al. (2012) provide evidence that mutual fund managers tend to reduce price impact, during the last financial crisis.[68] In this paper we follow a more general approach to explore which stocks are preferred to be sold, using a cross-sectional approach. Our results confirm that liquid stocks have been under much more selling pressure while in the same time gain a premium relative to illiquid stocks, during the financial crisis.

Lou and Sadka (2011) test the relation of stock returns with illiquidity level and illiquidity risk and show that illiquidity does not have explanatory power over the stock returns during the crisis. On the contrary, they find that illiquidity risk predicts the cross-sectional variation of stock returns during the same period. In our paper we use risk and turnover adjusted returns, in contrary to raw returns, that there were used by Lou and Sadka. We also explore in detail the mechanism behind the fail of illiquidity to predict stock returns in this period.

Florackis et al. (2014), using data from the London Stock Exchange, find that macro-liquidity shocks are transmitted in different ways to the stocks, depending on their illiquidity. They also find that liquid stocks are mainly affected by the macro-liquidity shocks, during the financial crisis of 2007-2009. A possible explanation of this phenomenon is the flight-from-liquidity, as investors leave the stock market by mainly selling liquid stocks to invest in more liquid asset classes (i.e. Treasuries). Nagel (2012) uses the returns of a short-term reversal strategy as a proxy for the gains of liquidity provision during turmoil periods. One of his results is that both liquid and illiquid stocks offer high expected returns for liquidity provision during a falling market, which is in line with our result regarding the fact that both liquid and illiquid stocks suffer large losses during the crisis.

Cella, Ellul and Giannetti (2013) find that during market turmoil episodes, short term institutional investors sell their holdings to a larger extent compared to long term institutional investors. As a result, stocks which are held mostly by short term institutional investors suffer greater losses due to intense price pressures. They also focus on the event of Lehman Brothers collapse, thus we see our paper as complementary, with our approach focusing on the role of liquidity during a market turmoil episode.

Our paper is also complementary to the strand of the literature that explores the relationship between illiquidity and stock returns, especially after a market shock (or generally when market illiquidity increases). Amihud (2002) finds that when expected market illiquidity rises (as it happens after a large market decline) there is a substitution effect from less liquid to more liquid stocks (flight to liquidity). This effect is translated as a premium of

[68] Manconi et al. (2010) report that during the financial crisis of 2007-2009, mutual funds decided to sell liquid bonds first.

liquid stocks in relation to illiquid ones. Amihud explores a large time span sample which imposes a difficulty to separate big shocks (that lead to funding problems of the investors) from "regular" increases of expected market illiquidity. We focus on a period of a severe crisis, during which funding constraints force investors to liquidate part of their positions. Thus we are able to identify two different opposite channels that emerge only after a large market decline (flight-to-liquidity and flight-from-liquidity) and subsequently control for the abnormal trading activity with the use of STF.

Acharya and Pedersen (2005) find that illiquidity level is cross-sectionally correlated with illiquidity risk, an indication of flight-to-liquidity. Like Amihud (2002), they also use a large time span. Their result of the high levels of cross-sectional correlation between illiquidity level and illiquidity risk is interesting and motivates us to also explore the relation between illiquidity risk and stock returns, using our measurement of cumulative abnormal returns. Our results indicate that illiquidity risk also affects the cross-section of stock returns during crisis with the same pattern that illiquidity level does. This fact emerges as a consequence from the high cross-sectional correlation between illiquidity level and illiquidity risk, questioning the usefulness of the latter for risk management purposes.

Our work is also related with the literature that focuses on the asymmetric relation of liquidity effect on asset prices. More specifically, there is evidence that only the sell-order illiquidity commands a premium, whereas the buy-order illiquidity does not appear to be priced. Brennan, Chordia, Subrahmanyam and Tong (2012) test the relation between illiquidity and asset pricing, using buy-order and sell-order illiquidity measures. They show that liquidity premium mainly emanates from the sell-order illiquidity. Brennan, Huh and Subrahmanyam (2013) test the relation between liquidity and stock returns, using a variant of the original Amihud illiquidity measure, namely the ratio of the absolute price change to the turnover, instead of the dollar volume.[69] They decompose their measure into elements that correspond to positive and negative return days and they show that only the element of the negative return days commands a return premium. They further analyze the positive and negative return elements using order flows and show that a sidedness variable accounting for sell order clustering on negative return days is associated with a larger part of the liquidity premium than the other liquidity components.

Although investors perceive liquidity as an insurance against large price declines, we show that this is not the case. A possible explanation regarding the seemingly failure of

[69] Floracki s et al. (2011) also use this variant of Amihud illiquidity measure, to control for any size bias.

liquidity to protect investors when market declines considerably, is that the measurement of liquidity in normal times does not adequately account for the possibility of correlated trades for funding liquidity reasons during a large market decline. Under these conditions, market participants may face correlated funding needs and "run" to the market (in analogous fashion as it is described by Bernardo and Welch (2004) and Morris and Shin (2004)) to absorb funding liquidity by mainly selling the low price impact liquid stocks. On aggregate the supply of liquid stocks dramatically increases, "converting" liquid stocks into illiquid, even if the cost per trade of them remains lower than that of the illiquid stocks.

Furthermore, our paper is related to the theoretical strand of the literature that investigates the limits of arbitrage. During a normal period, arbitrageurs exploit profitable opportunities by providing liquidity and correct any mispricing of the market. A necessary condition for arbitrageurs to provide their "services" and maintain the functionality of the markets is the availability of ample amounts of capital. The theoretical models of limits of arbitrage describe a number of different mechanisms that create funding problems to arbitrageurs, preventing them from providing liquidity to the market and eventually turn them into liquidity seekers.[70]

Two main theoretical predictions of the limits of arbitrage literature are relative to our empirical questions. The first states that investors prefer to hold liquid stocks after a negative shock to the market (flight-to-liquidity) to avoid suffering greater losses in case they need to sell their holdings in a subsequent period. Empirical literature confirmed this prediction. The second prediction of limits of arbitrage literature is that liquidity providers turn to liquidity seekers after a large market decline. However, what has not yet been analyzed is the choice of assets that constrained arbitrageurs will sell to meet their funding needs. Empirical papers that test the behavior of hedge funds and mutual funds during the crisis show that fund managers consider the price impact of their trades. The incentives of managers for choosing the asset that they sell to absorb liquidity is an issue that limits of arbitrage literature does not give a lot of attention. Our paper contributes as a simultaneous empirical test of both the aforementioned predictions of this literature. The novel result of the simultaneous existence of

[70] Limits of arbitrage literature is very extended. Brunnermeir and Pedersen (2009), Gromb and Vayanos (2002) and Anshuman and Viswanathan (2005) study the effect of binding margin constraints that force arbitrageurs to liquidate their investments. Vayanos (2004) and Garleanu and Pedersen (2007) study the effect of internal risk control rules. Shleifer and Vishny (1997), He and Krishnamurthy (2008) and Vayanos and Woolley (2008) explore the effect to the arbitrageurs in the presence of agency problems. Bernardo and Welch (2004) and Morris and Shin (2004) develop models of coordination failures, when a market-run (equivalent to the bank-run) occurs. Brunnermeir and Pedersen (2005) and Attari, Mello and Ruckes (2005) study the effect of predatory trading, occurring when a number of market participants exploit profits by trading against constrained market participants.

both flight-to-liquidity and flight-from-liquidity calls for further theoretical and empirical examination of the behavior of arbitrageurs after a negative market shock.

In addition, our work is connected with the branch of the literature that studies the effect of net order flows and order imbalances to stock returns. We include the STF to the standard four factor model under the assumption that order flows / order imbalances affect stock prices and stock returns. The fundamental reason for the interconnection between prices and order flows may be asymmetric information considerations (Kyle (1985)) or inventory management considerations (Ho and Stoll (1983), Spiegel and Subrahmanyam (1995)). Empirical studies confirm the significant contemporaneous relationship between net order flows (order imbalances) and returns (Stoll (2000), Chordia, Roll and Subrahmanyam (2002), Chordia and Subrahmanyam (2004)).

Blume, MacKinlay and Terker (1989) examine the relationship between order imbalances and stock price movements the days around the crash of 1987. They show that there is a strong relation between order imbalances and stock price movements both in time-series and cross-section analyses. They explain the larger price decline of stocks that were included in the S&P 500 relative to stocks that were not included on it, through the cross-sectional differences of their order imbalances. The subsequent price rebound the day after the crash, is also connected with order imbalances. Their analysis is analogous to the one in our paper, as they show that a part of the losses on S&P stocks at the day of the 1987 crash is related to abnormal selling pressure and not to real economic factors.

Finally, our paper is also related to the strand of the literature which studies the fact of downward slopping demand curves of stocks. Asset pricing models assume horizontal demand curves for stocks, where the price is an unbiased predictor of the real underlying value of the firm, a function of the future cash flows and of the discount factor. According to this assumption any excess demand for stock should not cause any (or almost any) change in the price of a stock. The papers of this literature show that in reality this is not the case, as the demand curves for stocks slope down (some papers of this literature is that of Shleifer (1986), Lynch and Mendenhall (1997) and Lou (2012) for stocks and that of Garleanu, Pedersen and Poteshman (2009) for the option pricing). The STF captures this non-fundamental price pressure.

3. Empirical Methodology

The purpose of this paper is to study the relation between illiquidity and cumulative abnormal stock returns during a crisis. We propose and implement a novel way for estimating abnormal returns, accounting for the impact of the abnormal turnover during a crisis and split abnormal returns into two parts: (a) the one associated with the trading for funding needs and (b) the other related with the need for reduction of portfolio illiquidity.

3.1 Signed Turnover Factor (STF)

To capture the part of the stock return that stems from non-fundamental factors we use a measure of the net demand for stocks. Order imbalances are the most sophisticated measure of this type. Order imbalance is measured as the difference between the buy and sell orders. Conventionally, when buy (sell) orders exceed sell (buy) orders, order imbalance has positive (negative) sign. The intuition is that if buy (sell) orders are more, it is more probable that the initiators of the trades are buyers (sellers), indicating the corresponding direction of the net demand for a stock. The estimation of order imbalances, however, requires a lot of intra-day data that are not available in a broad set of markets. Another measure that is commonly used is fund flows. This measure could be easily obtained and interpreted as pressure for buying or selling stocks. However, this measure inevitably contains information for the investing behavior of mutual or hedge funds and it not a clear net demand measure for the stock market.

We construct and employ to this study a low frequency (weekly) net stock demand measure, called signed turnover factor (STF). We do this in three steps:

i. We assign the sign of the daily return, to the number of shares traded on the same day.

ii. We sum the daily signed number of shares on a weekly basis.

iii. Finally, we divide the above summation with total number of shares, to obtain the weighted signed turnover.

$$STF_t = \frac{\sum_{d=1}^{5} sign(return)_d * (\#\ of\ shares\ traded)_d}{total\ \#\ of\ shares} \quad (1)$$

STF_t is the signed turnover factor for the week t, $sign(return)_d$ is the sign of the return of day d and $(\# \, of \, shares \, traded)_d$ is the number of shares traded at day d.

Our main assumption for the construction and use of this measure is the same that is made by Pastor and Stambaugh (2003)[71], that the volume of a day with the sign of the return of this day is a rough proxy of the order flow for a stock. We use the turnover instead of the volume to control for the size of the companies and have a clear measure of trading activity, comparable among the cross-section of the stocks. Another advantage of using the turnover instead of the volume is that the series becomes detrended, stationary and consequently suitable to be augmented as an extra factor in a time-series econometric specification with the returns as the dependent variable.

Although STF counts the whole turnover as the net order flow in a daily manner, whereas a part of it may be split between buy and sell orders. This could be misleading, as the information we need is the net demand for stocks and not the whole trading activity.[72] We consider that to a certain extent this problem is alleviated due to the weekly averaging of the daily signed turnovers.

An additional advantage of the use of weekly frequency is the exclusion of short-lived trading effects that may arise in a daily frequency. We thus focus on mid-frequency net demand effects that may have more permanent price impacts on the stocks. This way, the effect of the STF on returns is separated to some extend from the cost-of-trade liquidity effects.

3.2 Normal level of STF

STF is a variable that takes both positive and negative values, depending on the trading direction of each week. To define a normal **level** of STF, we use its absolute values.[73] Although STF does not exhibit any time-trend, its absolute values increase steadily through the years of our sample. That is, the volatility of STF increases through time. We thus estimate, for each stock, the mean of the absolute value of STF of the weeks t-52 to t, after adjusting for a time trend:

[71] "The basic idea is that, if signed volume is viewed roughly as "order flow",…", Pastor and Stambaugh, Journal of Political Economy, 2003, vol. 111, no. 3

[72] For a critique of the use of whole volume (or turnover) instead of the order imbalance you could see Chordia et al. (2002) and Chordia and Subrahmanyam (2004).

[73] The estimation of the mean or the median of STF would give us a very small value, not representative of a normal level of trading activity. Besides we want the information about the magnitude of the trading activity and not about the direction of the trading pressure.

$$(normal\ level\ of\ STF)_{i,t} = a_{i,t}^{STF} + b_{i,t}^{week} \cdot t \quad (2)$$

The mean absolute value of STF ($a_{i,t}^{STF}$) and the coefficient of the time trend ($b_{i,t}^{week}$) are retrieved from rolling time series regressions (with a window from t-52 to t-1) of the absolute value of STF for stock i on the number of the weeks:

$$(|STF|)_{i,t} = a_{i}^{STF} + b_{i}^{week} \cdot t + e_{i,t}^{STF} \quad (3)$$

We use this value as a benchmark of regular trading activity of a stock. We hence treat higher absolute values of trading activity as abnormal. Considering that we use cumulative abnormal values our inference is conservative, as we sum the excess turnover over several consecutive weeks. In unreported results we use the median and the mean of STF as the normal level of STF and the results do not considerably change.

We determine the normal level of STF using the trend-adjusted mean, due to the lack of a model that describes trading activity. Lo and Wang (2000, 2006) identify the absence of such models and propose a general intertemporal capital asset pricing model with simultaneous determination of prices and turnovers. Their empirical implementation does not include specific determinants of turnover, while it is beyond the scope of this paper to explore possible factors that may determine turnover. Dennis and Strickland (2002) use the median turnover from a prior period as the normal level of turnover.

3.3 The assumed model for the returns

We assume that stock returns are determined by a set of risk factors (excess market returns ($R_m - r_f$), the returns of a trading portfolio that is long in small stocks and short in big stocks (SMB), the returns of a trading portfolio that is long in high book to market stocks and short in low book to market stocks (HML), the returns of a portfolio that is long in stocks with high momentum (measured as the cumulative return of a stock for a period starting one year ago until three months before time t) and sort in stocks with low momentum (MOM)) plus a factor constructed by us (STF_i), individual for each stock, which accounting for the demand for stocks due to non-fundamental reasons:

$$r_{i,t} = a_i + b_i^{m} \cdot (R_m - r_f)_t + b_i^{SMB} \cdot (SMB)_t + b_i^{HML} \cdot (HML)_t + b_i^{MOM} \cdot (MOM)_t + b_i^{STF} \cdot STF_{i,t} + e_{i,t} \quad (4)$$

,where b_i^m, b_i^{SMB}, b_i^{HML} and b_i^{MOM} are the risk factor coefficients for the stock i, b_i^{STF} is the elasticity of the stock return for each unit of $STF_{i,t}$ for the stock i and $STF_{i,t}$ is the signed turnover factor for the stock i.

We estimate each stock's betas with the risk factors and our turnover factor using weekly returns from the beginning of 2003 until the end of May of 2008. We choose to use market-oriented risk factors to control effectively for contemporaneous changes in market valuation. In addition, by using these risk factors we are in position to isolate the effect of liquidity from other effects that may be correlated with illiquidity levels.

3.4 The event

We conduct an event study around the bankruptcy of Lehman Brothers at September of 2008 (15 weeks before and 37 weeks after the event). This event was not totally unexpected but it was a shock to the stock market and triggered a large decline to the main indexes.[74] Also the VIX index rose dramatically, as it is reported in Cella, Ellul and Giannetti (2013), who use the spikes of the VIX as a certain criterion of a negative shock to the market. We also conduct our analysis for the period between May 2004 and May 2009 to confirm that the effects we describe emerge during large market declines and are not constantly present.

3.5 The estimation of abnormal returns

For each week, we measure the abnormal part of the returns on a 26-week window, as the difference between the predictions of the 5-factor model and the realized returns. The gross returns can be decomposed as:

$$r_{i,t} = \hat{b}_i^m \cdot \left(R_m - r_f\right)_t + \hat{b}_i^{SMB} \cdot SMB_t + \hat{b}_i^{HML} \cdot HML_t + \hat{b}_i^{MOM} \cdot MOM_t + \hat{b}_i^{STF} \cdot (norSTF)_{i,t} + AR_{i,t} \quad (5)$$

, where $norSTF_i$ is the normal level of the trading activity factor for the stock i. We estimate this by giving to the pre-estimated normal level of STF the sign of the STF of the specific week, within the event window:

$$(norSTF)_{i,t} = (normal\ level\ of\ STF)_i \cdot sign\left(r_{i,t}\right) \quad (6)$$

[74] See Brunnermeier (2009) and Gorton (2008) for detailed "diaries" of the financial crisis of the period 2007-2009.

, where $sign(r_{i,t})$ is the sign of the STF of week t. For example, if for the stock "X" we estimate that the normal level of STF has the value 0.05, and at the first week of the event window the stock has negative STF, the value of the $norSTF$ for this stock for the first week is -0.05. If for the second week the STF is positive, the value of $norSTF$ is 0.05, and so on.

With $AR_{i,t}$ we symbolize the abnormal returns, which we estimate as the difference between the realized returns and the prediction of the model that we use:

$$AR_{i,t} = r_{i,t} - \left[\hat{b}_i^m \cdot \left(R_m - r_f\right)_t + \hat{b}_i^{SMB} \cdot SMB_t + \hat{b}_i^{HML} \cdot HML_t + \hat{b}_i^{MOM} \cdot MOM_t + \hat{b}_i^{STF} \cdot (norSTF)_{i,t}\right] \quad (7)$$

In our specification $AR_{i,t}$ includes two distinct parts:

$$AR_{i,t} = FFL_{i,t} + FTL_{i,t} \quad (8)$$

, where $FFL_{i,t} = \hat{b}_i^{STF} \cdot (abnSTF)_{i,t}$ (9) and $FTL_{i,t} = AR_{i,t} - \hat{b}_i^{STF} \cdot (abnSTF)_{i,t}$ (10)

The $FFL_{i,t}$ is the part of the abnormal return attributable to the flight from liquidity phenomenon. This component can be estimated directly, as the product between the pre-estimated coefficient of STF and the $abnSTF$, which is the difference between the realized STF and the $norSTF$. The $FTL_{i,t}$ is the part of the abnormal return attributable to the flight to liquidity phenomenon. The latter component can be estimated only indirectly, as the difference of $AR_{i,t}$ with $\hat{b}_i^{STF} \cdot (abnSTF)_{i,t}$.

We then compute the cumulative abnormal returns ($CAR_{i,t}$) as the sum of the abnormal returns in the relevant event window.[75] We use the $CAR_{i,t+26}$ (window of 26 weeks) for our analysis, and we also compute the cumulative values of the two parts of the $AR_{i,t}$, $CFFL_{i,t}$ and $CFTL_{i,t}$ for the same window.

3.6 The cross-sectional regressions

We retrieve the illiquidity premium as the coefficient of a cross-sectional regression of the cumulative abnormal returns of the window after week t on the illiquidity level (ILLIQ, Amihud (2002)) as it is measured at week t. The first specification that we use is a univariate one (which we also use for the whole period analysis), with the ln(ILLIQ) as the only explanatory variable:

$$CAR_{i,t+26} = a + \gamma \cdot (\ln (ILLIQ))_{i,t} + e_{i,t+26} \quad (11)$$

[75] Following Coval and Stafford (2007) and Cella et al. (2013).

We run regression (11) for each week starting from the first week of June of 2004 until May of 2009.

We retrieve FFL and FTL parts of the illiquidity premium as the coefficients of the cross-sectional regressions of CFFL and CFTL of the window after week t on the illiquidity level as it is measured at week t:

$$CFFL_{i,t+26} = a + \gamma \cdot (\ln(ILLIQ))_{i,t} + e_{i,t+26} \quad (11b)$$

and

$$CFTL_{i,t+26} = a + \gamma \cdot (\ln(ILLIQ))_{i,t} + e_{i,t+26} \quad (11c)$$

We also run the regressions (11b) and (11c) for each week starting from the first week of the June of 2004 until the May of 2009.

We also run multivariate regressions of the $CAR_{i,t+26}$, $CFFL_{i,t+26}$ and $CFTL_{i,t+26}$ of 19 and 26 weeks after the collapse of Lehman Brothers, on the illiquidity measure ILLIQ and on a number of other control variables (including an illiquidity risk measure) as they are measured 15 weeks before the event:

$$CAR_{i,t+26} = a + \gamma \cdot (\ln(ILLIQ))_{i,t} + B' \cdot Controls_i + e_{i,t+26} \quad (12)$$

$$CFFL_{i,t+26} = a + \gamma \cdot (\ln(ILLIQ))_{i,t} + B' \cdot Controls_i + e_{i,t+26} \quad (12b)$$

$$CFTL_{i,t+26} = a + \gamma \cdot (\ln(ILLIQ))_{i,t} + B' \cdot Controls_i + e_{i,t+26} \quad (12)$$

where B' is a vector with the coefficients of each one of the control variables that we use.

4. Data

We use data from New York Stock Exchange (NYSE) stocks, which we obtain from Bloomberg. These data include stock prices, gross returns (price change plus dividend yield), market capitalization, market-to-book value ratio, share volume and total number of shares. We obtain the above data both in daily and in weekly frequencies to estimate some weekly variables by the daily observations of the corresponding week.

We obtain data for the market excess return, the two additional Fama-French risk factors and the momentum factor (the latter only in daily frequency) by the Kenneth French's website. Since there is no weekly momentum available in the site, we construct the weekly momentum risk factor from our sample. We construct 769 different time series of the momentum factor, excluding each time one firm of our sample to avoid mechanical correlation. We use for each firm its corresponding momentum factor that we construct. For the construction of the weekly momentum factor we form two portfolios, one with stocks with the higher returns for the prior year and one with stocks with the lower returns for the prior year. Each week we readjust the portfolios, but because we measure annual returns, the adjustments are not substantial. The momentum factor is the return of a portfolio long to the portfolio of the past winners and short to the portfolio of the past losers. We also obtain from Kenneth French's website the risk-free rate.

We obtain data on the quarterly holdings of institutional investors from Thomson One. These data are from 13F mandatory institutional reports which are filed with the Securities and Exchange Commission (SEC). The 13F form requires all the institutions that have discretion of over $100 million at the end of the calendar year, to report their long holdings in the next year.

We use only common stocks that are traded in New York Stock Exchange (NYSE). We choose not to include stocks from AMEX, because they have smaller capitalization and higher asymmetric information problems. We also do not include NASDAQ stocks in our sample because of the differences on the reporting of trading volume.[76] We exclude ADRs, REITs, preferred stocks and other publicly traded investment instruments to avoid including in the study unusual characteristics. The initial sample size was 1430 companies. We obtain data for the period between January 2003 and May 2009. We exclude companies with at least one of the following characteristics:

i. Companies with missing values on the data.
ii. Companies with negative values in the market-to-book ratio.
iii. Companies with price more than $1000.
iv. Financial or real estate companies. We exclude these two categories of companies from our sample because they were in the epicenter of the crisis and asymmetric information issues as well as credit risk issues were emerged.

After the exclusion of companies that meet the abovementioned criteria, the sample is reduced to 769 companies. The remaining firms are distributed well in terms of market capitalization, liquidity and industry sectors.

[76] NASDAQ is a dealer market and the share volume is measured double. NYSE is an auction market.

5. Variables

In this section we describe the independent variables that we use in the econometric analysis. All the variables that we use (dependent, independent and auxiliary) are described in the Appendix A, on the Table A1.

5.1 The illiquidity measure: ILLIQ

The basic independent variable of our analysis is ILLIQ (Amihud, 2002):

$$ILLIQ_{i,t} = \frac{1}{Days} \cdot \sum_{d=1}^{Days} \frac{|R_{i,d}|}{V_{i,d}} \quad (13)$$

For each stock i and for each week t we measure the $ILLIQ_{i,t}$ for a period of one year prior to t. $Days$ is the number of daily observations we use for the estimation (the total number of trading days of the period between t-52 to t). For the econometric analysis regarding the period after Lehman's event we choose to use the illiquidity value of 3 months before the event to avoid endogeneity between the measure and the stock returns[77]. $|R_{i,d}|$ is the absolute value of the return of day d and $V_{i,d}$ is the dollar volume of day d, for stock i.

The intuition of $ILLIQ$ is that a stock is relatively illiquid when its price moves a lot in response to low volume. We assume that $ILLIQ$ incorporates all the trading costs, such as broker fees, bid-ask spreads, market impact and search costs. Besides, $ILLIQ$ is widely used as an illiquidity measure by the vast majority of the relative literature and in addition is the low frequency measure with the highest correlation with high-frequency measures of illiquidity (Goyenko et al. (2009), Hasbrouck (2009)).

The main advantage of $ILLIQ$ over bid-ask spreads is that it does not depend on the size of the trade, while spreads refer to a specified relatively small number of shares. In any case, we also repeat our analysis with the bid-ask spread as a percentage of the stock price, to have a comparable result among the cross-section. The results are almost unaffected by the change in the illiquidity measure, a fact that we expect since the cross-sectional correlation between the illiquidity measures is very high ($corr(ILLIQ, spread) = 0.87$).

[77] Hameed et al. (2010) show that illiquidity increases after a market decline.

5.2 Illiquidity risk

We include the illiquidity risk as a control variable to our multivariate examination. Illiquidity risk is defined as the covariation of stock returns with marketwide liquidity. Stocks which commove with systematic marketwide liquidity have higher expected returns, because they do not protect investors from illiquidity when they mostly need it. By its definition it is a measure that should predict stock returns during a large market decline. However there is empirical evidence that liquidity per se and liquidity risk are distributed in similar fashion in the cross-section of the stock market (Acharya and Pedersen, 2005). For this reason it is attractive to study the dynamics of illiquidity risk and whether it is in position to predict stock returns during the crisis, despite the high cross-sectional correlation with illiquidity level. We also use illiquidity risk as the single independent variable on the cross-sectional regressions.

We measure illiquidity risk by running the rolling time-series regressions for each stock, with a window of 52 weeks:

$$(r_{i,t} - r_{f,t}) = a_i + b_{m,i} \cdot (R_m - r_f)_t + b_{illiq,i} \cdot (MILLIQ)_t + e_{i,t} \quad (14)$$

,where b_i^{illiq} is the illiquidity risk beta, representing the magnitude of illiquidity risk of stock i, $(MILLIQ)_t$ is the innovation of marketwide illiquidity at time t. We construct the marketwide illiquidity factor as the equally weighted mean of the individual $ILLIQ$ of our sample's stocks. We measure the innovations of the marketwide illiquidity factor as the residuals from an AR(2) specification, following the relative literature (Acharya and Pedersen, 2005)[78].

5.3 Other control variables

In the multivariate cross-sectional regressions we also use a number of additional variables, as controls. We use the market beta and the beta of the STF, which we obtain from the estimation of the risk factor model of the stock returns (of equation (5)). The market beta is a basic variable that affects stock returns. Stocks with higher beta may exhibit worse abnormal returns, because they carry more market risk. We use STF beta to be sure that the effects that we find are not due to any mechanical reason due to the model we use to explain the stock returns. We also use the percentage of the ownership that is held by institutional holders, to ensure that our results are general and not valid only for institutional investors. We

[78] Illiquidity is persistent.

also use the mean turnover to ensure that our estimation about the abnormal STF is not driven mechanically by a higher general level of turnover of a stock.

We also use the market-to-book ratio to control for any value-growth considerations. Momentum is another variable that may affect the abnormal stock returns, as investors may wish to realize the gains of the previous period. We include the momentum of each stock, measured for the period from 1/6/2007 until 31/5/2008, as the cumulative return. We also use the ratio of debt-to-assets to control for the firms' leverage. Finally we use the return on assets (ROA) and the return on equity (ROE) as additional control variables. All the variables are *ex ante* measured, synchronized with the measurement of $ILLIQ$. The accounting variables are taken from the last report before the 31/5/2008.

6. Validation of STF

We continue with a preliminary analysis of the nature of STF. We first show at Table 1 some descriptive statistics of the marketwide STF, which is the time-series of the cross-sectional mean of individual STFs. The mean STF over the whole sample (Panel A) is 0.06%, meaning that on average there was a buying pressure over the period between 2003 and 2009, including the years of the financial crisis. The standard deviation and the mean of the absolute value of STF is considerably higher, 3.16% and 1.77% respectively. The absolute value of STF is the base for the estimation of the normal level of STF. Panel B of Table 1 shows the same statistics, only using the "normal" period. Mean STF is 50% higher (on average 0.09%) if we exclude the period after Lehman Brothers event. The standard deviation and the mean absolute value of STF are slightly lower compared to that of the full sample. The mean STF of the event period is -0.15%, indicating the high selling pressure of that period. The standard deviation and the mean absolute value of STF during the event period are much higher from that of the "normal" period, showing that not only the direction of trade changed to selling, but the intensity of trade changed as well, exhibiting a considerable increase.

We proceed with the validation of STF as a proxy of the net non-fundamental demand for stocks. Table 2 illustrates the correlation of the marketwide STF with the percentage ETF flows, the weekly changes of Libor-OIS spread, the weekly changes of VIX, and the 4 factors of the Carhart model. The first column shows the statistics estimated over the whole sample period, the second column shows the statistics of the "normal" period and the third column shows the statistics of the event period.

The correlation of marketwide STF with the ETF flows over the whole period is low (0.07). However, their correlation is significantly higher (0.17) if we measure it only during the "normal" period, validating the relation of STF with the demand for ETFs. Their correlation turns negative (-0.09) during the event period, probably because investors turn to ETFs during the market fall, in a meaning of flight-to-quality. Moreover, the correlation of the excess market return with the ETF flows is also negative when we measure it during the event period.

The correlation of marketwide STF with the Libor-OIS spread is negligible during the "normal" period (-0.03), but it is negative and significant (-0.25) during the event period, indicating that there is relation between funding illiquidity and STF. When there are funding constraints, investors sell stocks to retrieve cash. The correlation with the VIX changes is also negative and significant, independently of the period of the estimation. An increase in VIX is contemporaneously related with selling pressures and vice versa.

Marketwide STF is highly correlated with contemporaneous excess market returns (0.91), an expected result. Their correlation remains in similar levels even when we use only the "normal" or the event period. The high correlation between marketwide STF and excess market returns is an indication that at least to an extent the former is a proxy for fundamental demand as well. Marketwide STF is also correlated with SMB factor (0.47, "normal" period), but the correlation drops to 0.16 in the event period. It seems that there is a disconnection between the "size" premium and the trading activity, during the event. An even more intriguing relation is revealed between the marketwide STF and the HML factor. While in "normal" period they are uncorrelated (0.02), their correlation during the event jumped to 0.61. Taking into account that during the event the mean STF is negative, we detect a relation between the selling pressure and the reverse of the value premium. A similarly impressive change in the correlations before and after the event is derived for marketwide STF and MOM factor (0.14 at "normal" period and -0.68 at event period), a result that is in agreement with the "momentum crash" described in Daniel and Moskowitz (2016).

The high correlation between marketwide STF and excess market returns raises concerns about the validity to use STF together with market excess returns in the same time-series regressions on individual stock level. We thus estimate the time-series correlations of the individual STFs with each of the 4 factors and then estimate their cross-sectional means and medians. The results are reported in Table 3. As Table 3 shows, the average (and the median) correlation of the individual STF with the marker excess return is not very high (0.37

mean and 0.38 median), while the mean correlations of individual STF with the other three factors are even smaller. It is not, hence, redundant to augment the 4-factor model with STF.

Finally, we want to check whether there is a mechanical relation between stock returns and STF, since we construct the latter by taking the signs of the daily stocks returns during each week. We thus measure the percentage of the firm-week pair observations in which the sign of STF is the opposite of the sign of the return. Almost in 21% of the total number of firm-week observations (251,817), the sign of STF is the opposite from the return sign. This evidence alleviates the concerns about a pure mechanical correlation between STF and returns.

Table 1: Descriptive statistics of STF and abs(STF)

The table provides descriptive statistics over the pooled sample, for the whole period (Jan2003 – May2009, 335 weeks), for the "normal" period (Jan2003 – May2008, 283 weeks) and for the event period (Jun2008 – May2009, 52 weeks). The mean, standard deviation, skewness, kurtosis, minimum, median and maximum values are reported for the STF and the absolute value of STF (abs(STF)).

	mean	s.d.	skewness	kurtosis	min	median	max
Panel A: Whole Period (Jan2003 – May2009)							
STF	0.06%	3.16%	-2.13	101.42	-187.53%	0.07%	71.44%
abs(STF)	1.77%	2.62%	7.41	191.00	0.00%	0.98%	187.53%
Panel B: "Normal" Period (Jan2003 – May2008)							
STF	0.09%	2.83%	-2.70	150.20	-187.53%	0.08%	70.28%
abs(STF)	1.58%	2.35%	8.54	290.13	0.00%	0.88%	187.53%
Panel C: Event Period (Jun2008 – May2009)							
STF	-0.15%	4.52%	-0.97	29.98	-102.60%	-0.02%	71.44%
abs(STF)	2.74%	3.60%	4.98	55.99	~0.00%	1.64%	102.60%

Table 2: The correlations of the marketwide STF with ETF flows, Libor-OIS, VIX, and the 4 Factors of Carhart model, estimated over the whole sample period (Jan 2003 – May 2009), over the "normal" period (Jan 2009 – May 2008) and over the event period (Jun 2008 – May 2009)

Time-series contemporaneous correlation between marketwide STF (the time-series of the cross-sectional mean of STF) and a number of systemic factors. ETF-flows is the aggregate flows to ETFs as a percentage of their total assets, Δ(Libor-OIS) is the change of the Libor-OIS from week t-1 to week t, Δ(VIX) is the change of the VIX index from week t-1 to week t, Rm-Rf is the excess return of the market, SMB is the return of the factor small-minus-big, HML is the return of the factor small-minus-big and MOM is the return of the factor winners-minus-losers. Data for STF, ETF-flows, Libor-OIS, VIX and MOM are from Bloomberg. Data for Rm-Rf, SMB and HML are from Kenneth French's site (which also includes their definitions). MOM is constructed according the definition provided in the site of Kenneth French, with data from our sample, because this factor is not provided in weekly frequency.

	Jan 2003 – May 2009	Jan 2003 – May 2008	Jun 2008 – May 2009
ETF-flows	0.07	0.17	-0.09
Δ(Libor-OIS)	-0.16	-0.03	-0.25
Δ(VIX)	-0.71	-0.69	-0.74
Rm-Rf	0.91	0.88	0.94
SMB	0.34	0.47	0.16
HML	0.38	0.02	0.61
MOM	-0.29	0.14	-0.68

Table 3: The cross-sectional mean and median of the time-series contemporaneous correlation of the individual STFs with the 4 Factors of Carhart model, estimated over the whole sample period (Jan 2003 – May 2009).

Cross-sectional mean and median of the contemporaneous time-series correlation between individual STFs and the 4 Factors of Carhart model. For each firm of the sample, we esteimate the time-series correlation of its STF with the 4 factors and then we estimate the cross-sectional mean and median of these correlations. Rm-Rf is the excess return of the market, SMB is the return of the factor small-minus-big, HML is the return of the factor small-minus-big and MOM is the return of the factor winners-minus-losers. Data for STF, and MOM are from Bloomberg. Data for Rm-Rf, SMB and HML are from Kenneth French's site (which also includes their definitions). MOM is constructed according the definition provided in the site of Kenneth French, with data from our sample, because this factor is not provided in weekly frequency.

	Rm-Rf	SMB	HML	MOM
mean correlation	0.37	0.13	0.15	-0.12
median correlation	0.38	0.13	0.15	-0.11

7. Whole period analysis

We proceed with a preliminary analysis of the relation between ILLIQ and stock returns during the period between June 2004 and May 2009 (257 weeks). For each week (t) from the first week of the July 2004 until the last week of May 2009 we run three separate cross-sectional regressions (769 stocks):

1. The cumulative risk and turnover adjusted stock returns of the last 26 weeks ($CAR_{i,26}$) on the ILLIQ value 26 weeks before, at (t-27), which is measured from the week (t-78) to the week (t-27).

$$CAR_{i,t+26} = a + \gamma \cdot (\ln(ILLIQ))_{i,t} + e_{i,t+26} \quad (11)$$

2. The cumulative FFL part of the returns of the last 26 weeks on the ILLIQ value 26 weeks before, at (t-27), which is measured from the week (t-78) to the week (t-27).

$$CFFL_{i,t+26} = a + \gamma_{FFL} \cdot (\ln(ILLIQ))_{i,t} + e_{i,t+26} \quad (11a)$$

3. The cumulative FTL part of the returns of the last 26 weeks on the ILLIQ value 26 weeks before, at (t-27), which is measured from the week (t-78) to the week (t-27).

$$CFTL_{i,t+26} = a + \gamma_{FTL} \cdot (\ln(ILLIQ))_{i,t} + e_{i,t+26} \quad (11b)$$

Figure 1 shows that the γ coefficient (illiquidity premium) of the effect of ILLIQ on CAR (cumulative abnormal returns) remains the same independently of which factor model we use (either a standard 4-factor model or the 5-factor model with the augmented STF). We are thus confident that our model is at least as good as the standard one and at the same time it enables us to split the abnormal stock returns into the two aforementioned parts. (Figure 1.B in Appendix B shows the t-statistics of the corresponding coefficients of the two models).

Figure 2 shows the relation between the γ coefficient (illiquidity premium) and the cumulative returns of NYSE composite index over a period of 26 weeks. As Figure 3 shows, within the crisis illiquidity premium increased, which means that illiquid stocks perform better than the liquid ones, the opposite of the flight-to-liquidity prediction. Especially after Lehman's event, the manifestation of this (opposite of the theory) phenomenon is striking. Only several weeks after the outburst of the event the illiquidity premium starts to fall eventually becoming negative. However when market rebounds, γ remains negative indicating that liquid stocks perform better than illiquid ones, while we would expect the opposite, according to the flight-to-liquidity hypothesis and the V-shape of the liquidity-related price impact. Overall, this evidence contradicts the flight-to-liquidity hypothesis and supports a flight-from-liquidity story.

Figure 3 illustrates the γ_{FFL} coefficient (from equation 11.A) and the cumulative abnormal STF (CASTF) over 26 weeks, an index of the cumulative abnormal trading pressure. It is clear that when CASTF becomes negative, after Lehman, the γ_{FFL} increases significantly, indicating that the negative relative performance of liquid stocks over the illiquid ones, is connected with the increased selling pressure of the crisis. This result confirms the hypothesis, that during a crisis, a phenomenon of flight-from-liquidity emerges according to which investors prefer to sell liquid stocks to absorb funding liquidity. Even if liquid stocks have low transaction costs, the fact that there is high selling pressure on them results in a large aggregate negative price impact. This is the first study that shows the existence of the flight-from-liquidity phenomenon in a cross-section of stocks.

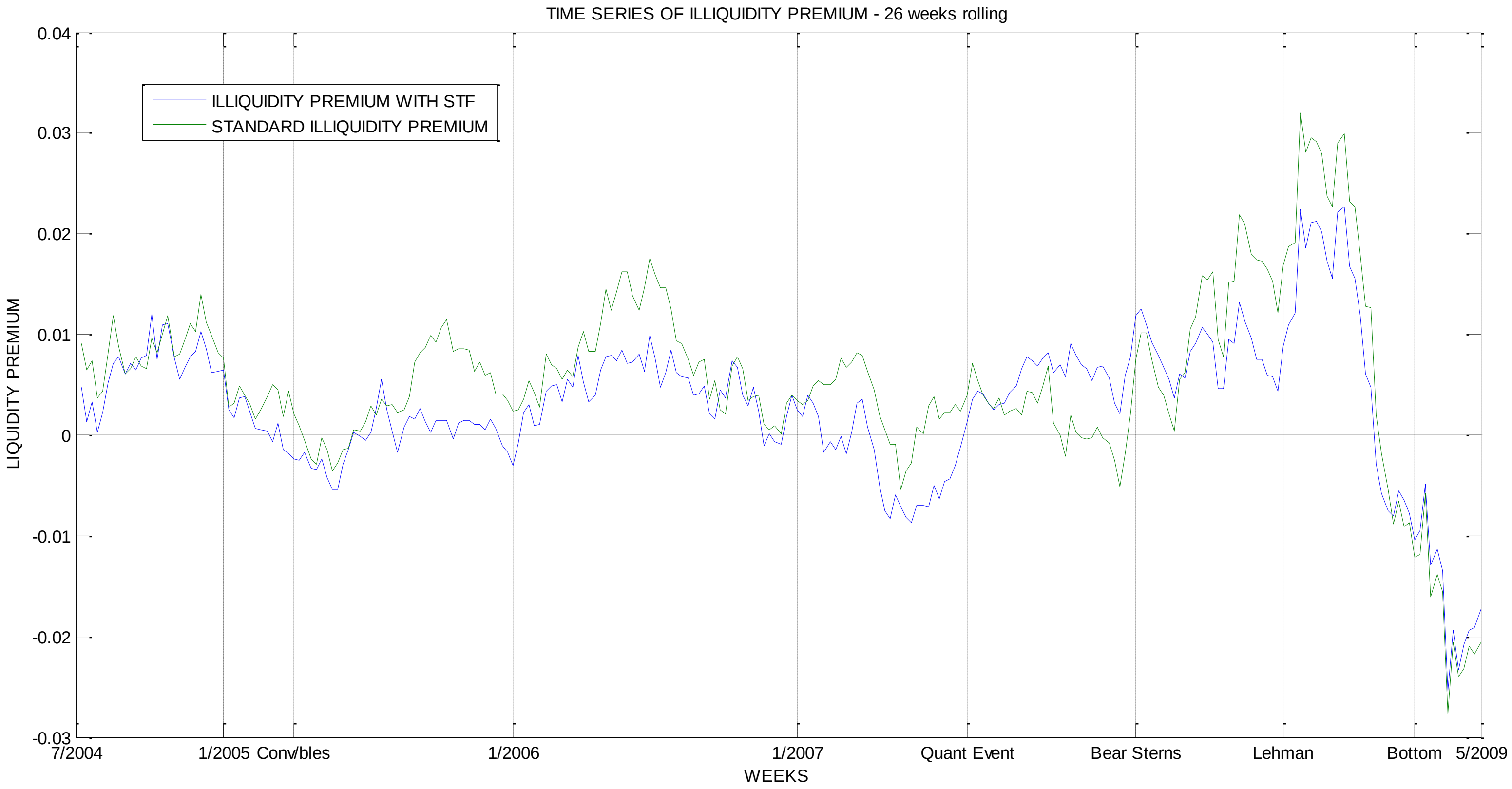
Figures 3.B and 3.C (in Appendix B) show that in the normal period the γ_{FFL} is small and not statistically significant, further supporting our hypothesis. γ_{FFL} increases and becomes strongly significant when the crisis deepens, after the Bear Sterns event and especially after the Lehman event (Figure 3.D in Appendix B). Finally, it is very clear that when the market rebounds, this coefficient become negative and statistical significant, a rather expected result, as investors step in and buy again the liquid stocks in relatively low prices. Besides, the V-shape of returns is a standard feature, when illiquidity shocks occur. What is new evidence is that this occurs to the liquid stocks. The behavior of this coefficient during all the sample period strongly supports the hypothesis of flight-from-liquidity.

Figure 4 illustrates the γ and its two parts the γ_{FFL} and γ_{FTL}. It is clear in this figure that the variation of the FFL part of the illiquidity premium is rather too small to explain the time-variation of the whole illiquidity premium, until the outburst of the crisis. On the other hand, the FTL part of the illiquidity premium seems to be the driver of the time-variation of the illiquidity premium, until the outburst of the crisis. Figures 5 and 6 show that after the outburst of the crisis, the FFL part is the one driving the illiquidity premium, instead of the FTL part which is disconnected from the whole illiquidity premium.

Finally, Figures 7 and 7.B – 7.C (in the Appendix B) show the FTL part of the illiquidity premium. This is the part of the illiquidity premium which shows the “pricing” of liquidity after considering the level (and the direction) of trading (essentially it is the illiquidity premium after the adjustment for the amortizing cost of trade). Hence, FTL is the part which includes any flight-to-liquidity manifestation in stock returns, since this effect is a “pricing” effect. A negative γ_{FTL} means that the prices of liquid stocks further increased relative to the prices of the illiquid ones, and it is thus interpreted as indication of a flight-to-liquidity phenomenon. Our results show that flight-to-liquidity incidents occur in more than one instances (and especially during the convertible bonds crisis, at the start of 2007 and some weeks after Lehman collapse).[79] It also seems that after controlling for the level of trading activity (FFL part), the flight-to-liquidity phenomenon is revealed, confirming the predictions of the related literature.

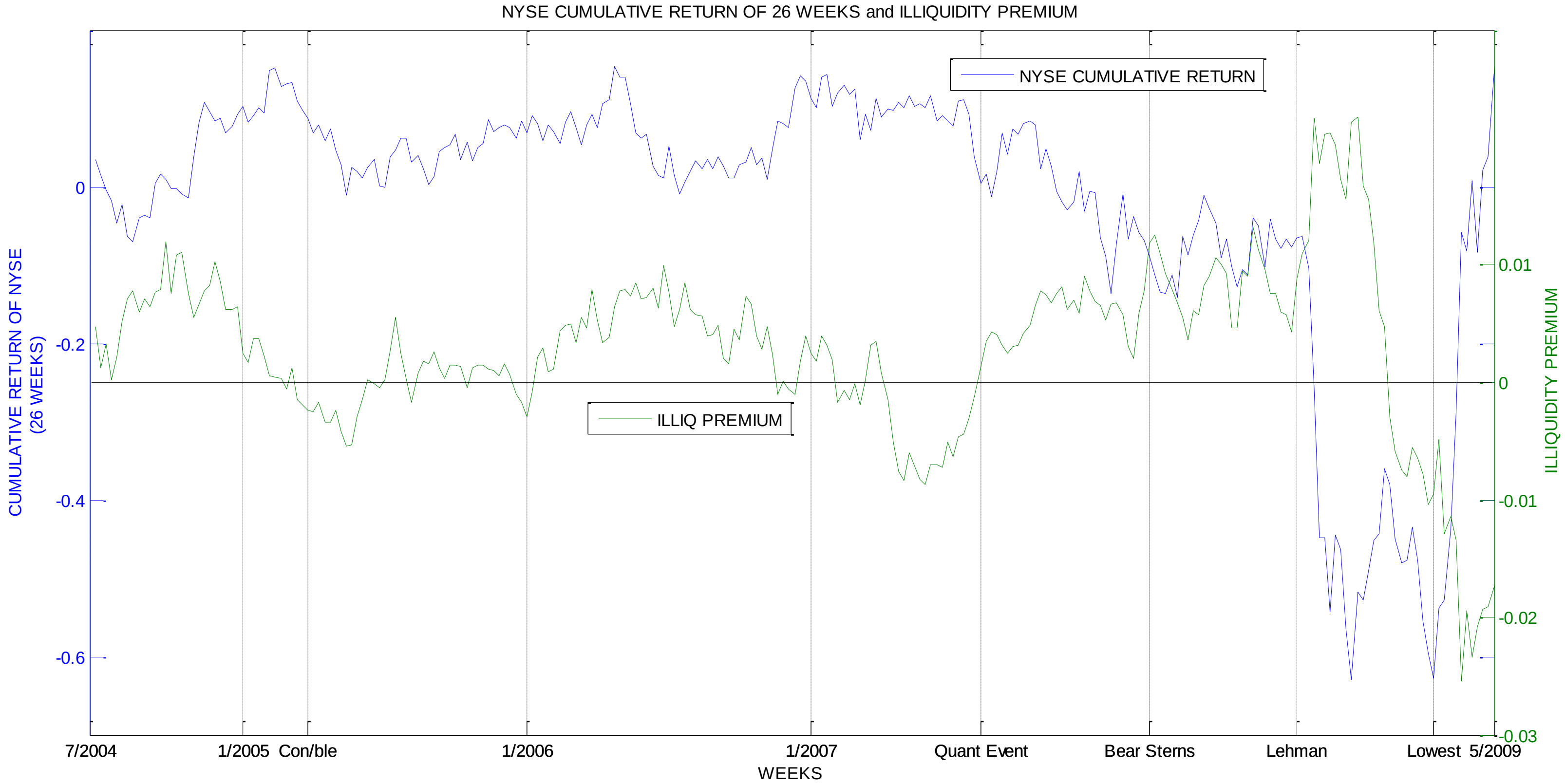
[79] A very interesting by-product of our analysis, is that FTL part is negative in certain periods, even if the whole premium is positive.

Figure 1: Illiquidity premium (γ coefficient): Classical approach Vs STF augmented approach.



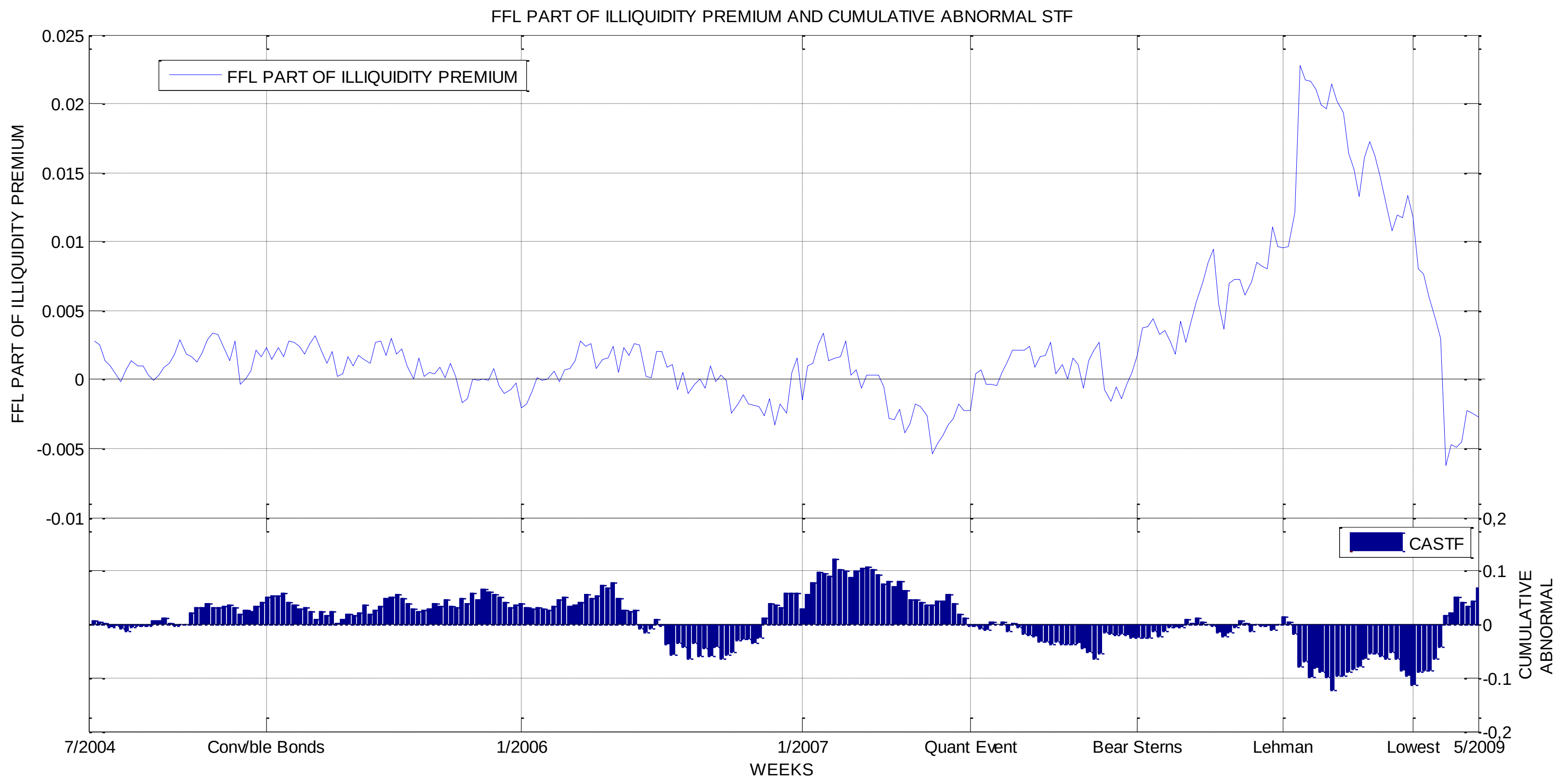
The figure illustrates the evolution of the illiquidity premium which is estimated using both the classical approach (green line) and STF augmented approach (blue line), over the last 284 weeks of the sample (July 2004 to May 2009). Under the classical approach, the illiquidity premium is the coefficient (γ) of rolling cross-sectional regressions of the risk-adjusted returns (4-factor model) of each week on the ILLIQ of the 52 previous weeks. Under our approach, the illiquidity premium is the coefficient (γ) of rolling cross-sectional regressions of the risk and turnover adjusted returns (4-factor model plus STF) of each week on the ILLIQ of the 52 previous weeks.

Figure 2: The illiquidity premium (γ coefficient) (green line, right Y-axis) and the cumulative 26-week returns of NYSE composite index (blue line, left Y-axis).



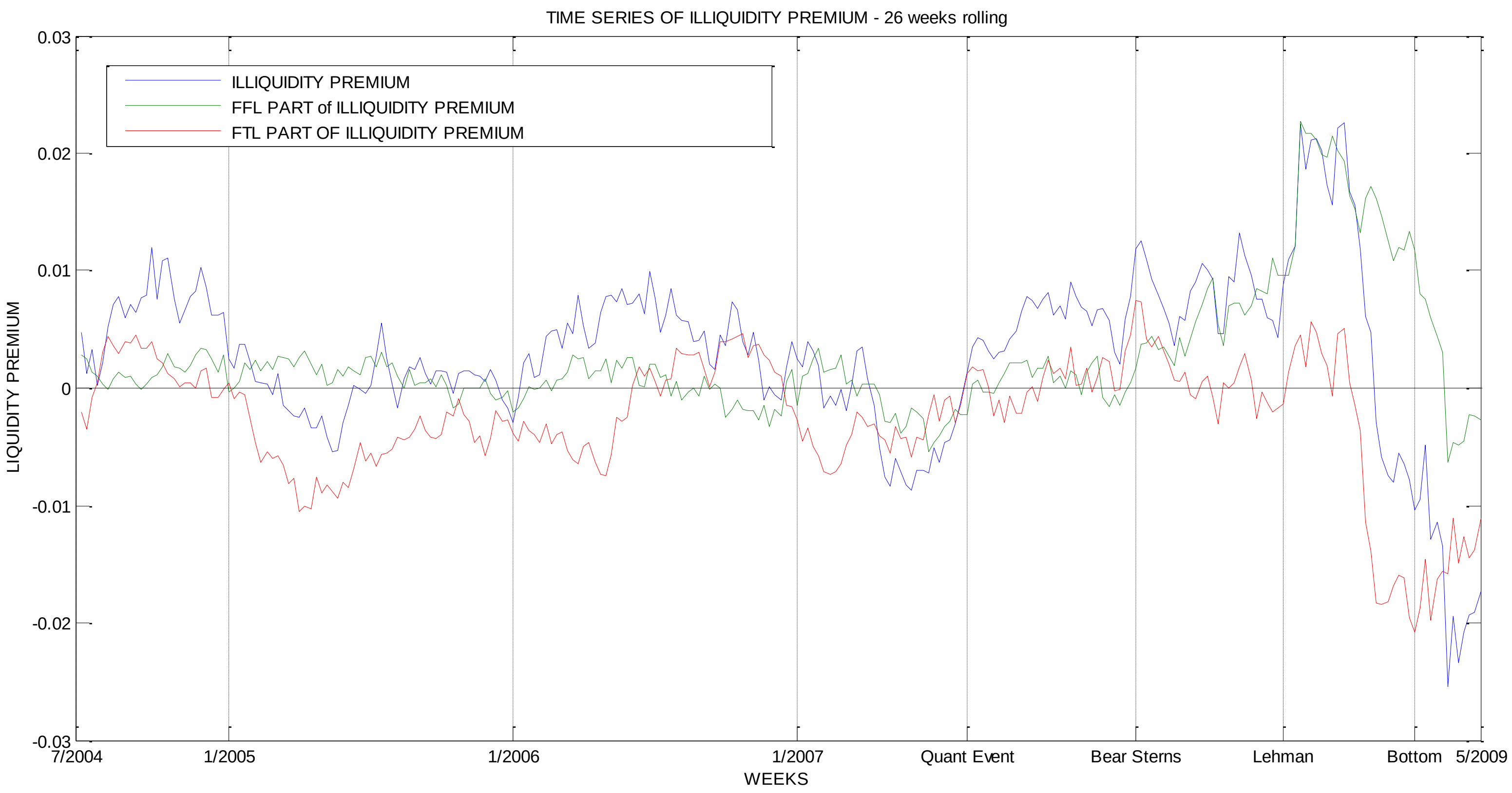
The figure illustrates the rolling illiquidity premium (γ coefficient) (green line, right Y-Axis) and the rolling cumulative 26-week return of NYSE composite index, over the last 257 weeks of the sample (July 2004 to May 2009). The illiquidity premium is the coefficient (γ) of rolling cross-sectional regressions of the risk and turnover adjusted returns (4-factor model plus STF) of each week on the ILLIQ of the 52 previous weeks.

Figure 3: The FFL part of the illiquidity premium (γ_{FFL} coefficient) (blue line, upper part of the figure) and the cumulative abnormal STF of 26 weeks (bold blue bars, lower part of the figure).



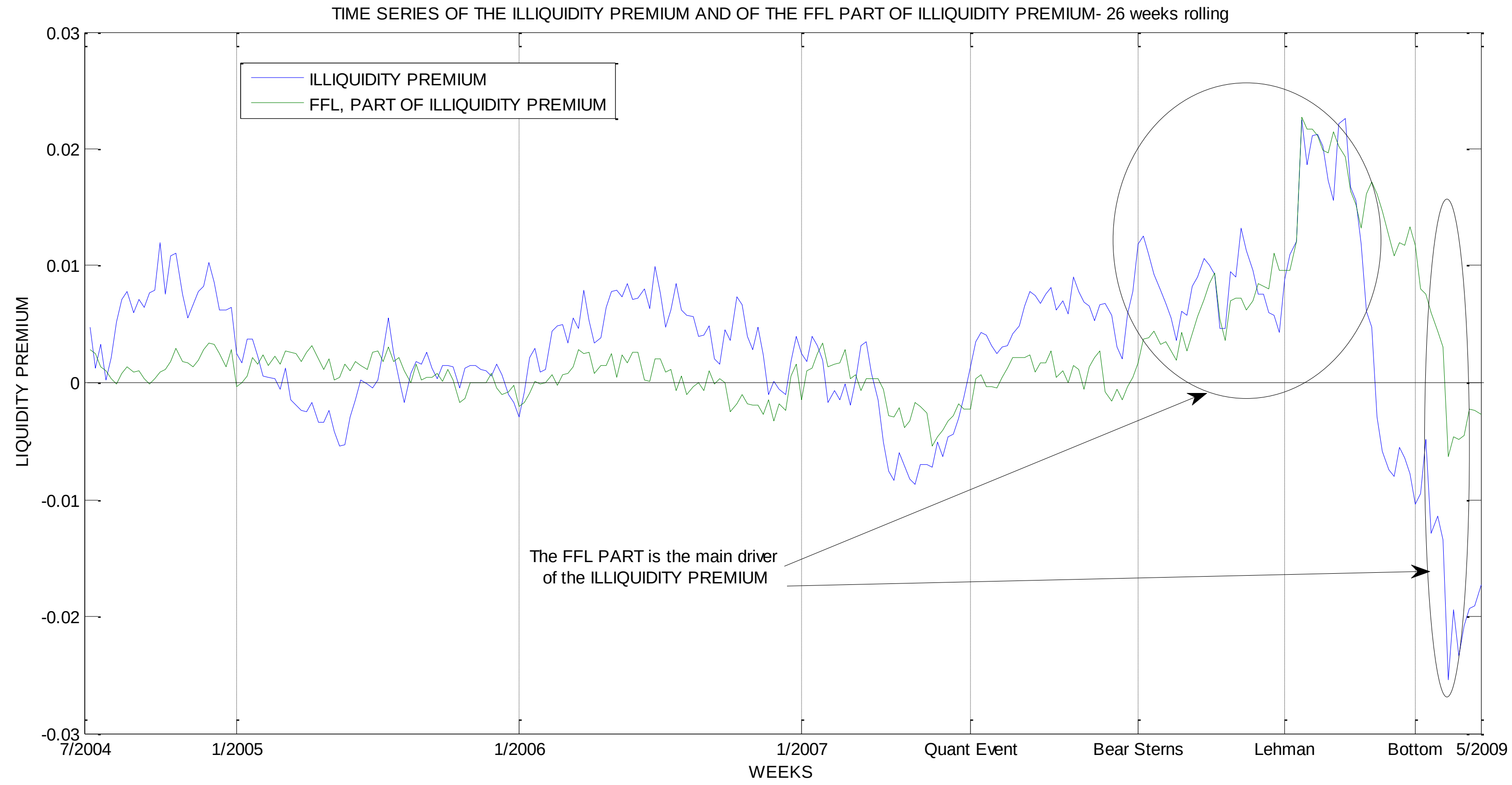
The upper part of the figure illustrates the rolling FFL part of the illiquidity premium (coefficient γ_{FFL}), over the last 257 weeks of the sample (July 2004 to May 2009). The FFL part of the illiquidity premium is the coefficient (γ_{FFL}) of rolling cross-sectional regressions of the abnormal STF-related returns of each week on the ILLIQ of the 52 previous weeks. The abnormal STF-related returns are the product of the abnormal STF with the pre-estimated elasticity of stock returns on STF (b_{stf}). The lower part of the figure illustrates the rolling cumulative abnormal STF (CASTF) over a period of 26 weeks. A positive value of CASTF indicates a net buying pressure over the last 26 weeks.

Figure 4: The illiquidity premium (γ coefficient) (blue line), the FFL part of the illiquidity premium (γ_{FFL} coefficient, green line) and the FTL part of the premium (γ_{FTL} coefficient, red line).



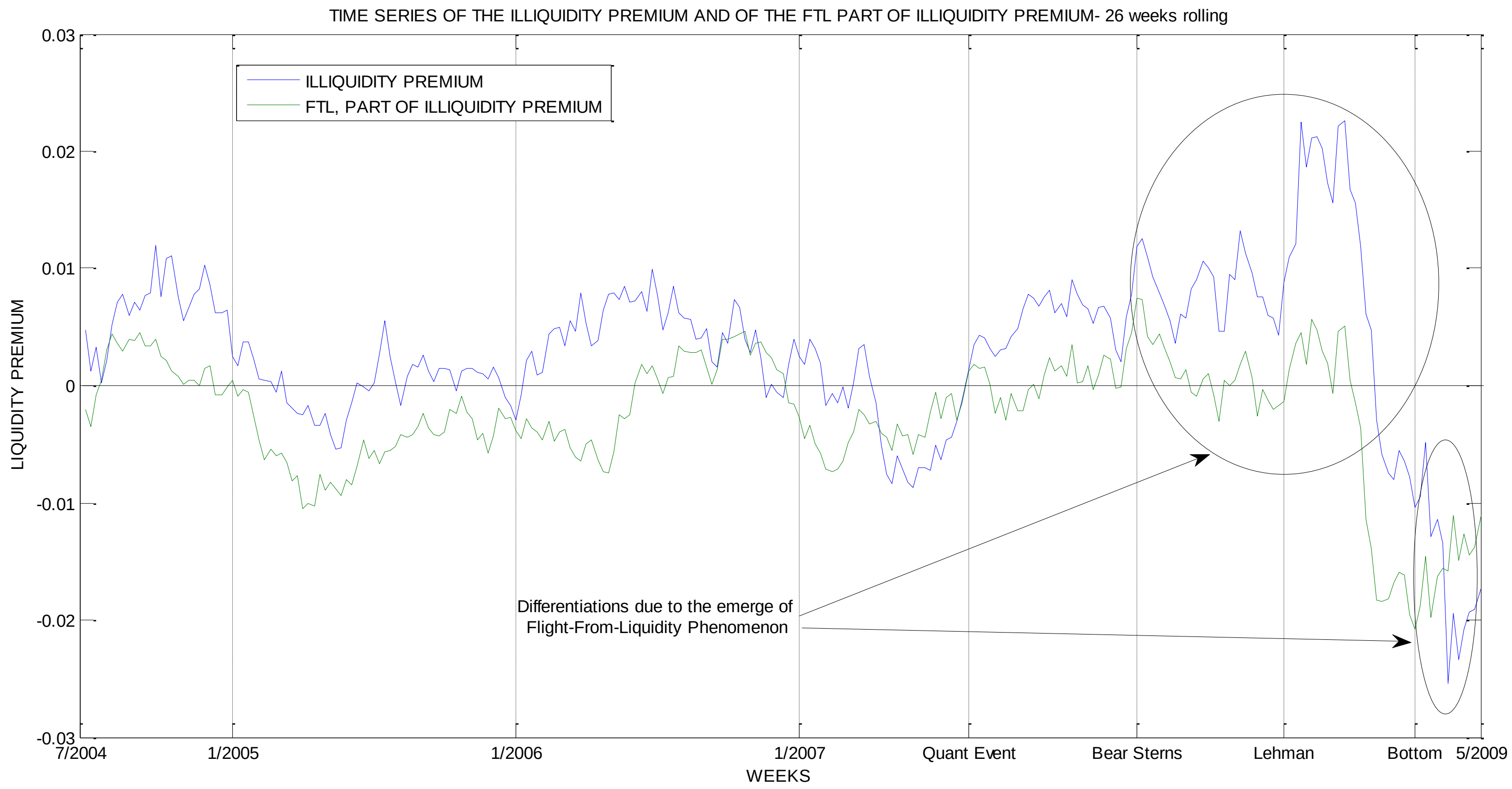
The figure illustrates the rolling illiquidity premium (γ coefficient) (blue line), and its two parts, the FFL (green line) and the FTL (red line). The illiquidity premium is the coefficient (γ) of rolling cross-sectional regressions of the risk and turnover adjusted returns (4-factor model plus STF) of each week on the ILLIQ of the 52 previous weeks. The FFL part of the illiquidity premium is the coefficient (γ_{FFL}) of rolling cross-sectional regressions of the abnormal STF-related returns of each week on the ILLIQ of the 52 previous weeks. The abnormal STF-related returns are the product of the abnormal STF with the pre-estimated elasticity of stock returns on STF (b_{stf}). The FTL part of the illiquidity premium is the coefficient (γ_{FTL}) of rolling cross-sectional regressions of the risk and turnover adjusted returns (after the subtraction of the abnormal STF-related returns) of each week on the ILLIQ of the 52 previous weeks.

Figure 5: The whole illiquidity premium (blue line) and the FFL part of the illiquidity premium (green line).



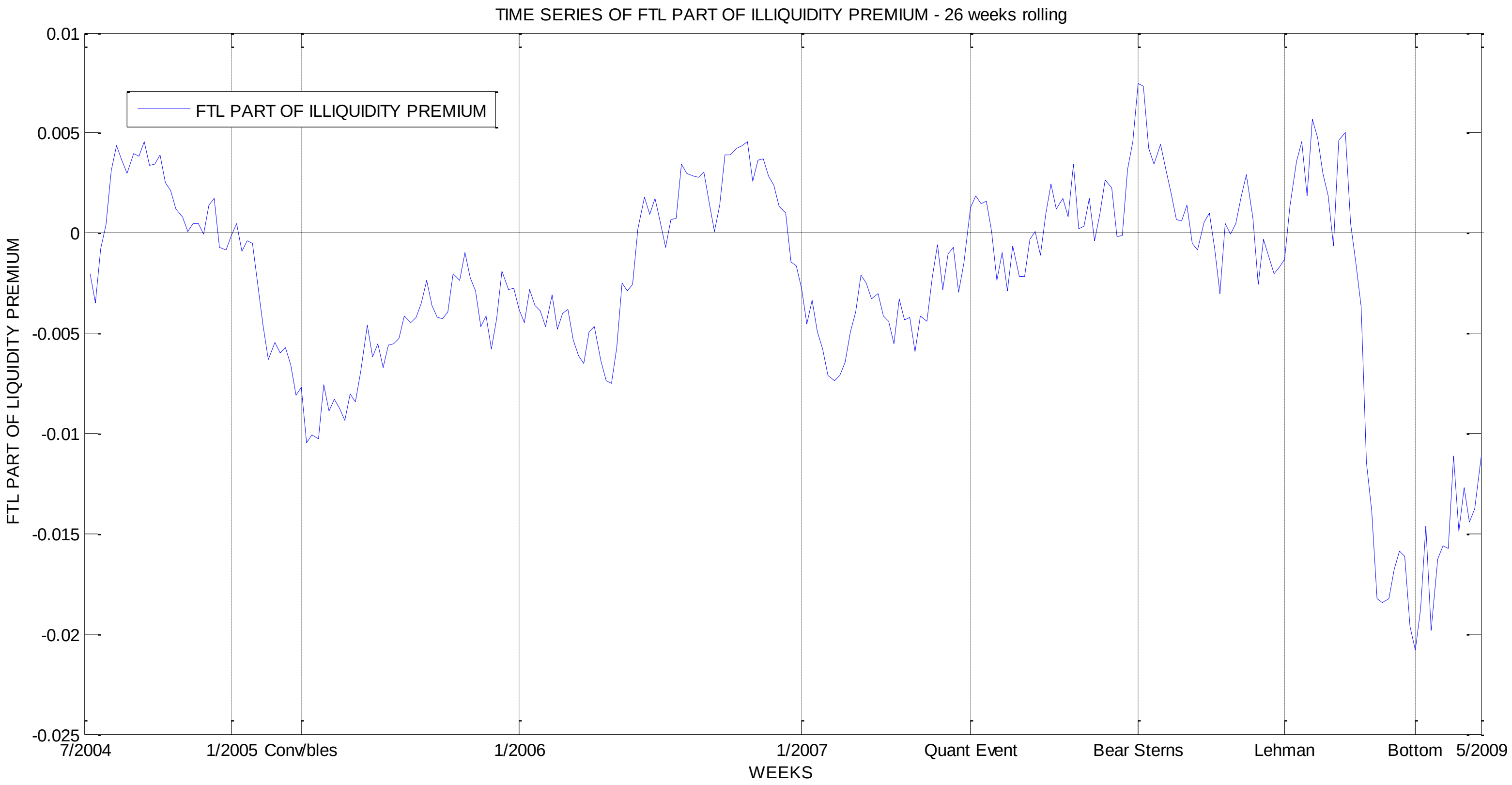
The figure illustrates the rolling illiquidity premium (γ coefficient) (blue line) and the FFL part of the illiquidity premium (green line). The illiquidity premium is the coefficient (γ) of rolling cross-sectional regressions of the risk and turnover adjusted returns (4-factor model plus STF) of each week on the ILLIQ of the 52 previous weeks. The FFL part of the illiquidity premium is the coefficient (γ_{FFL}) of rolling cross-sectional regressions of the abnormal STF-related returns of each week on the ILLIQ of the 52 previous weeks. The abnormal STF-related returns are the product of the abnormal STF with the pre-estimated elasticity of stock returns on STF (b_{stf}). The two circles indicate two periods in which the main driver of illiquidity premium is its FFL part.

Figure 6: The whole illiquidity premium (blue line) and the FTL part of the illiquidity premium (green line).



The figure illustrates the rolling illiquidity premium (γ coefficient) (blue line) and the FTL part of the illiquidity premium (green line). The illiquidity premium is the coefficient (γ) of rolling cross-sectional regressions of the risk and turnover adjusted returns (4-factor model plus STF) of each week on the ILLIQ of the 52 previous weeks. The FTL part of the illiquidity premium is the coefficient (γ_{FTL}) of rolling cross-sectional regressions of the risk and turnover adjusted returns (after the subtraction of the abnormal STF-related returns) of each week on the ILLIQ of the 52 previous weeks. The two circles indicate two periods in which illiquidity premium is disconnected from its FTL part.

Figure 7: The FTL part of the illiquidity premium (coefficient γ_{FTL}).



The figure illustrates the rolling FTL part of the illiquidity premium (coefficient γ_{FTL}), over the last 257 weeks of the sample (July 2004 to May 2009). The FTL part of the illiquidity premium is the coefficient (γ_{FTL}) of rolling cross-sectional regressions of the risk and turnover adjusted returns (after the subtraction of the abnormal STF-related returns) of each week on the ILLIQ of the 52 previous weeks.

8. Descriptive statistics

In this section we provide some descriptive statistics regarding the main variables that we use in the econometric analysis of the event period. Table 4 shows the main statistics of the three main dependent variables (CAR, CFFL and CFTL), and of the independent variables. The mean CAR at this point of time for our sample is almost -11%, from which almost the -7% is attributable to the CFFL part (trading) and the remaining ~-4% is attributable to the CFTL (the "pricing" effects).

Table 5 shows the cross-sectional correlation between the independent variables that we use in the econometric analysis of the event. ln(ILLIQ) and ln(spread) are highly correlated (0.878) as it is known for the various illiquidity measures. The correlation between ln(ILLIQ) and ln(size) it is very high in value and negative (-0.938), an expected result, since there is a mechanical relation between them. We thus do not include both ln(ILLIQ) and ln(size) simultaneously in the regressions. The correlation between ln(ILLIQ) and ln(volatility) is high and positive (0.457). As with ln(size) we avoid to include ln(volatility) to the same regressions with ln(ILLIQ). However, we run separate regressions using ln(size) and ln(volatility) as the main independent variables. In these regressions we also include the ln(turnover), which is also mechanically related with ln(ILLIQ), although their correlation is relatively low (-0.129).

Table 5 also reveals the high correlation between ln(ILLIQ) and illiquidity beta (-0.301). The sign of the coefficient is negative, because more negative illiquidity betas indicate higher illiquidity risk.[80] High correlation between illiquidity per se and illiquidity risk is an evidence of the related literature (Amihud (2002), Acharya and Pedersen (2005)). This correlation is interpreted as evidence of flight-to-liquidity, in the sense that an illiquid stock is the stock with the worse returns during a market downturn. However, according to our results (Section 8), illiquidity risk is also affected from the flight-from-liquidity phenomenon. It is exactly the high correlation of illiquidity risk with illiquidity per se that make the former a weak predictor of stock returns during large market declines.

The STF beta by its nature is an illiquidity measure of price impact, since it shows the change in the price for a given level of trading pressure. One concern that we have is how much it differs from the basic dependent variable of our econometric analyses. The cross-sectional correlation between ln(STF beta) and ln(ILLIQ) is 0.304, while the correlation of

[80] In a stock with negative market beta, an increase in market illiquidity induces a decrease in its price.

the former with ln(spread) is 0.250 (Table 5). We are thus confident that our findings are not imposed by us through a mechanical relation between the separation of the abnormal returns with STF, and the analysis with ln(ILLIQ) as main dependent variable.

The rest of the correlations of ln(ILLIQ) with the independent variables are the expected ones. Stocks with higher market beta, lower market-to-book ratio and lower roa and roe are more illiquid. The correlation between the two alternative illiquidity variables (ln(ILLIQ) and ln(spread)) is very high (0.878) as it is already known from the related literature. The correlations of ln(spread) with the rest of the independent variables are similar with that of ln(ILLIQ) with them.

Finally, the comparison between market beta and ln(STF beta) reveals that the two variables are adequately different in their nature. The correlation between them is insignificant (-0.048), even if the correlation of each of them with ln(ILLIQ) and ln(spread) is around 0.30. Market beta is highly correlated with both size and volatility, while STF beta it is not. On the other hand, the correlation of ln(STF beta) with ln(turnover) is -0.714 – and it is to a large extend mechanical – when the correlation of market beta with ln(turnover) is 0.202. Finally, the correlation of ln(STF beta) with illiquidity beta is negligible (0.045), but that of market beta with illiquidity beta is -0.374, indicating that STF beta is significantly different from market beta and illiquidity risk.

Table 4: Descriptive statistics of main variables.

The table provides descriptive statistics over the cross-section of the variables of the main econometric analysis. 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*
CAR_{19} (%)	-10.969	32.383	-0.349	4.385	-97.878	-8.187	68.763
$CFFL_{19}$ (%)	-7.155	38.333	-19.149	468.155	-58.429	-4.949	44.019
$CFTL_{19}$ (%)	-3.814	42.027	10.488	213.939	-84.308	-1.819	66.315
ln(ILLIQ)	-2.673	1.772	0.158	2.872	-6.631	-2.729	1.488
ln(spread)	-6.740	0.571	0.690	4.600	-7.919	-6.782	-5.013
ln(size)	21.822	1.544	0.342	2.768	18.884	21.701	25.690
ln(volatility)	-3.155	0.331	0.060	3.044	-3.903	-3.148	-2.431
ln(turnover)	-4.255	0.576	0.061	3.798	-5.736	-4.247	-2.899
illiquidity beta	-0.070	0.071	-0.565	4.717	-0.293	-0.063	0.098
ln(inst-perc)	-0.178	0.244	-7.378	114.140	-0.933	-0.127	0.000
market beta	0.601	0.232	0.693	3.963	0.159	0.574	1.278
ln(STF beta)	0.314	0.494	1.439	8.428	-0.604	0.247	2.125
ln(mtb)	1.062	0.604	0.834	5.310	-0.072	1.018	2.997
momentum	-0.234	0.259	-0.314	4.892	-0.996	-0.224	0.404
debt-to-assets	22.543	14.821	0.447	2.970	0.000	21.960	60.410
RoA (%)	5.082	7.409	-6.288	112.837	-12.770	4.852	22.882
RoE (%)	12.045	29.854	-12.601	307.308	-30.040	11.866	69.778

Table 5: Correlation matrix between the independent variables.

Correlation Matrix between the independent variables, which are used in the econometric analysis. The values of the independent variables are from the end of May 2008. The sample includes 769 stocks.

	ln (ILLIQ)	ln (spread)	ln(size)	ln (volatility)	ln (turnover)	illiquidity beta	ln(inst-perc)	market beta	ln (STF beta)	ln(mtb)	momentum	debt-to-assets	RoA	RoE
ln(ILLIQ)	1													
ln(spread)	0.878	1												
ln(size)	-0.938	-0.818	1											
ln(volatility)	0.457	0.575	-0.489	1										
ln(turnover)	-0.129	-0.027	-0.086	0.512	1									
illiquidity beta	-0.301	-0.321	0.360	-0.419	-0.220	1								
ln(inst-perc)	-0.066	-0.118	-0.014	0.119	0.306	-0.052	1							
market beta	0.311	0.336	-0.308	0.520	0.202	-0.374	0.127	1						
ln(STF beta)	0.304	0.250	-0.027	-0.042	-0.714	0.045	-0.250	-0.048	1					
ln(mtb)	-0.306	-0.251	0.306	0.024	0.038	0.191	-0.019	-0.147	0.011	1				
momentum	0.135	0.087	-0.159	-0.134	-0.016	-0.267	-0.053	-0.123	-0.025	-0.265	1			
debt-to-assets	-0.033	0.003	0.023	-0.120	-0.048	-0.025	-0.037	-0.042	-0.066	-0.061	0.035	1		
RoA	-0.200	-0.246	0.158	-0.191	0.056	0.012	0.426	-0.174	-0.144	0.145	0.120	-0.132	1	
RoE	-0.237	-0.291	0.202	-0.220	0.020	0.030	0.493	-0.124	-0.147	0.108	0.070	0.027	0.825	1

9. Econometric results: Event period

9.1 Cross-sectional regressions of CAR on ILLIQ and additional control variables.

We first report the results of the cross-sectional regression of $CAR_{i,t+26}$ (where t+26 refers to the 19th week after Lehman) on the $ln(ILLIQ)_{i,t}$ and on other regressors (Table 6). The results are indicative for the non-predictive ability of ln(ILLIQ) on cumulative abnormal stock returns during a large market decline. The coefficient of ln(ILLIQ) is not statistically significant in any of the specifications. In addition, neither illiquidity beta seems to have predictive ability over stock returns during the crisis. The adjusted R-squares are quite small, but this is expected as we use individual stocks on the regressions. Besides, our dependent variable is the abnormal returns during a crisis, which includes a lot of noise. The 7.90% adjusted R-square at the full specification (5), in which we include the full set of control variables, is the highest one.

At specification (3) to (5) we include the natural logarithm of the percentage of the institutional ownership on each stock. It seems that higher level of institutional ownership is related with lower abnormal returns after Lehman event. One possible reason is the fund redemptions that funds faced during that period, in addition to the generalized funding illiquidity problems. These constraints may lead a large majority of institutional investors to sell stocks to retrieve cash from the stock market. Market beta predicts higher abnormal returns during the event period; however this is an additional effect after the risk-adjustment. Higher share turnover and higher STF beta predicts lower abnormal returns, an indication that STF beta has a predictive ability during large market declines. Higher momentum predicts lower abnormal returns, confirming the findings of Daniel and Moskowitz (2016) about the momentum crashes. Finally, it seems that stocks with higher roa perform better during the event period, an indication of flight-to-quality.

We repeat the same analysis using the 26-week CAR, 26 weeks after the collapse of Lehman, when illiquidity premium falls considerably as it is shown in Figure 1. Table 6b in Appendix C reports the results. At the univariate case (column (1)) and when we include illiquidity beta as the only extra control (column (2)) the coefficient is higher (in absolute value) relative to the respective ones of Table 6, but still insignificant. However, when we include additional controls (columns (3) to (5)) the coefficient of ln(ILLIQ) becomes statistically significant, indicating the flight-to-liquidity phenomenon. Nevertheless, the results remain weak. The coefficients of the extra controls do not change considerably compared to that of the basic specifications of Table 6.

Table 6: Cumulative abnormal stock returns (CAR) after Lehman's collapse and pre-crisis ILLIQ

Cross-sectional Robust OLS regressions of the cumulative abnormal stock return (CAR) of stock i 19 weeks after the collapse of Lehman Brothers, $CAR_{i,t+26}$, on the natural logarithm of ILLIQ of stock i, $log(ILLIQ)_{i,t-35}$, as it is observed 15 weeks before the collapse of Lehman Brothers, and on other lagged control variables for stock i, $Z_{i,t-35}$, which are also observed 15 weeks before the collapse of Lehman Brothers:

$$CAR_{i,t+26} = \alpha + \lambda \cdot (\ln(ILLIQ))_{i,t-35} + \Gamma' \cdot Z_{i,t-35} + e_{i,t+26}$$

There are 5 regressions in columns 1 through 5. The variables of each regression are described in the very left column. See Table 1 for the detailed definitions of the variables.

CARs are measured in percentage form. The sample consists of 769 stocks. t-statistics are inside the parentheses below the regression coefficients. 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)
ln(ILLIQ)	**-0.29** **(-0.48)**	**-0.30** **(-0.49)**	**-0.45** **(-0.72)**	**-1.09** **(-1.57)**	**-0.76** **(-1.06)**
illiquidity beta (*100)	-	-2.04 (-0.13)	-6.81 (-0.44)	1.23 (0.07)	-29.34* (-1.67)
% of inst/nals	-	-	-11.52*** (-2.63)	-11.16** (-2.45)	-15.35*** (-2.85)
market beta	-	-	-	1.66** (2.26)	17.94*** (3.33)
ln(STF beta)	-	-	-	3.60 (1.52)	-9.48*** (-2.86)
ln(mean turnover)	-	-	-	-	-16.59*** (-5.72)
ln(market-to-book)	-	-	-	-	-0.37 (-0.19)
momentum	-	-	-	-	-10.73** (-2.35)
debt-to-assets	-	-	-	-	0.01 (0.20)
RoA	-	-	-	-	0.93*** (3.48)
RoE	-	-	-	-	-0.09 (-1.42)
Adj-R^2 (%)	~0.00%	~0.00%	0.95%	2.23%	7.90%

9.2 Cross-sectional regressions of CFFL and CFTL on ILLIQ and additional control variables.

We then focus our analysis to the two parts of the CAR, namely the $CFFL$ and the $CFTL$. The summation of these two parts gives us the CAR. The $CFFL_{i,t+26}$ is related with the abnormal trading activity during the crisis (thus mainly selling activity), while the $CFTL_{i,t+26}$ is related with the shift of investor preferences towards more liquid stocks during the crisis. For each of the two parts (measured the 19th week after Lehman collapse), we run cross-sectional regressions on ln(ILLIQ) and on other controls (the same controls that we use for the cross-sectional regression of CAR, to have comparable results). We illustrate the results of the cross-sectional regressions of CFFL on Table 7 and of CFTL on Table 8.

In every specification of Table 7 ln(ILLIQ) enters with a positive and statistical significant coefficient. We remind that ILLIQ measures illiquidity and thus a larger value of it means an illiquid stock. The positive sign of the ILLIQ coefficients means that illiquid stocks perform better than the liquid ones; regarding the part of the abnormal returns that is related to excess trade (liquid stocks lose more due to excess selling during the crisis). In addition it seems that ln(ILLIQ) is by far the more significant determinant of the excess selling, a reasonable result when considering the incentive of fund managers to reduce the cost of trading during the crisis. These results further confirm the existence of a flight-from-liquidity phenomenon that is shown in the analysis of the whole period.

Illiquidity beta enters the regressions with a negative coefficient, which is marginally significant only in the specification that is shown column (4). The negative sign of the illiquidity beta indicates that stocks with more negative values, thus higher illiquidity risk, perform better after the event, regarding the CFFL part of the abnormal returns. That is, both illiquidity per se and illiquidity risk are affected from the flight-from-liquidity, a rather expected result since illiquidity risk is cross-sectionally correlated with illiquidity. The empirical evidence of previous studies that use long time series interpret the high correlation between illiquidity risk and illiquidity per se as evidence of flight-to-liquidity. However, a closer look at the data of the severe crisis of 2007-2009, reveals that not only flight-to-liquidity is “hidden” from a contemporaneous flight-from-liquidity effect, but also illiquidity risk is not a good predictor of the returns, exactly due to its high correlation with illiquidity itself.

Table 7b in Appendix C reports the coefficients from the regressions of $CFFL_{i,t+26}$ at the 26th week after the collapse of Lehman (it is the corresponding of Table 6b). The coefficient of ln(ILLIQ) is again positive and statistically significant, while the coefficients of the rest of the variables are not strong.

Our results of the regressions of $CFTL_{i,t+26}$ (table 3) on ln(ILLIQ) indicate that the latter predicts lower cumulative abnormal returns that are related with the shift on the preferences of investors towards liquid stocks. The coefficient of the ln(ILLIQ) is negative and statistical significant in every specification we test, which means that less liquid stocks perform worse than the liquid ones, confirming the flight-to-liquidity prediction. Contrary to the CFFL case (Table 7), when we examine the CFTL (Table 8) a lot of control variables are also significant.

Illiquidity beta enters the regressions with a positive coefficient (except full specification, column (5)), but statistically insignificant, except in the specification of column (4). The positive sign of the illiquidity beta indicates that stocks that carry higher illiquidity risk have worse performance after the event, regarding the CFTL part. However, the statistical significance of the coefficient of illiquidity beta is not very strong. On the other hand, the coefficient of the institutional percentage of ownership is negative and strongly statistically significant, indicating that after the event stocks with higher level of institutional ownership performed worse in terms of CFTL. Higher market beta, lower STF beta, lower share turnover, lower momentum and higher roa predict higher CFTL part of the abnormal returns after the collapse of Lehman Brothers. The effects of the rest of the variables on the total CAR are mainly coming from the effects of them to CFTL.

We repeat our analysis using as dependent variable the 26-week CFTL of the 26th week after the bankruptcy of Lehman Brothers (Table 8b in Appendix C, it is the corresponding of Table 6b and Table 7b). The coefficient of ln(ILLIQ) remains become more negative and significant, a result that we expect by the inspection of the Figures 4 and 7. At the last phase of the large market decline after the Lehman's event, flight-to-liquidity became more severe. In this case, the coefficient of illiquidity beta is negative, but again only marginally significant in two of the four specifications that we include it. The coefficient of the percentage of institutional ownership is much smaller and only in one case (column (4)) statistically significant. The coefficients of the rest of the variables do not exhibit any significant change. At the later phase of the event study, market-to-book value and roa seem to be very significant.

Overall, the results of the cross-sectional regressions of the two parts of the cumulative abnormal returns strongly support the hypothesis of the simultaneous existence of two opposite liquidity effects to stock returns during a large market decline. The two effects offset each other, and as a result illiquidity does not have predictive ability over stock returns during the crisis. The simultaneous existence of flight-from-liquidity and flight-to-liquidity is in line with the results of Ben-David et al. (2011) who find that hedge funds sold during the crisis high volatility stocks (which are also the illiquid stocks, especially during a crisis) and liquid stocks. Our result is also in line with the result of Lou and Sadka (2011) and of Nagel (2012), that liquid stocks did not overperform illiquid stocks during the financial crisis of 2007-2009.

Table 7: Cumulative abnormal FFL part of returns (CFFL) after Lehman's collapse and pre-crisis ILLIQ

Cross-sectional Robust OLS regressions of the cumulative FFL part of abnormal stock return (CFFL) of stock i 19 weeks after the collapse of Lehman Brothers, $CAFFL_{i,26}$, on the natural logarithm of ILLIQ of stock i, $log(ILLIQ)_{i,t-35}$, as it is observed 15 weeks before the collapse of Lehman Brothers, and on other lagged control variables for stock i, $Z_{i,t-35}$, which are also observed 15 weeks before the collapse of Lehman Brothers:

$$CFFL_{i,t+26} = \alpha + \lambda \cdot (\ln (ILLIQ))_{i,t-35} + \Gamma' \cdot Z_{i,t-35} + e_{i,t+26}$$

There are 5 regressions in columns 1 through 5. The variables of each regression are described in the very left column. See Table 1 for the detailed definitions of the variables.

CFFLs are measured in percentage form. The sample consists of 769 stocks. t-statistics are inside the parentheses below the regression coefficients. 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)
ln(ILLIQ)	**1.62*** (5.82)**	**1.49*** (5.09)**	**1.50*** (5.10)**	**1.41*** (4.39)**	**1.39*** (4.12)**
illiquidity beta (*100)	-	-10.76 (-1.48)	-10.37 (-1.43)	-15.12** (-1.97)	-10.83 (-1.32)
% of inst/nals	-	-	0.57 (0.28)	1.73* (1.82)	0.88 (0.35)
market beta	-	-	-	-2.69 (-1.13)	0.00 (~0.00)
ln(STF beta)	-	-	-	2.20** (1.97)	0.80 (0.50)
ln(mean turnover)	-	-	-	-	-1.86 (-1.35)
ln(market-to-book)	-	-	-	-	-0.05 (-0.06)
momentum	-	-	-	-	4.28** (2.00)
debt-to-assets	-	-	-	-	-0.06* (-1.92)
RoA	-	-	-	-	0.12 (0.98)
RoE	-	-	-	-	-0.00 (-0.19)
Adj-R^2 (%)	0.02%	0.16%	0.04%	1.91%	4.75%

Table 8: Cumulative abnormal FTL part of returns (CFTL) after Lehman's collapse and pre-crisis ILLIQ

Cross-sectional Robust OLS regressions of the cumulative FTL part of abnormal stock return (CFTL) of stock i 19 weeks after the collapse of Lehman Brothers, $CAFTL_{i,t+26}$, on the natural logarithm of ILLIQ of stock i, $log(ILLIQ)_{i,t-35}$, as it is observed 15 weeks before the collapse of Lehman Brothers, and on other lagged control variables for stock i, $Z_{i,t-35}$, which are also observed 15 weeks before the collapse of Lehman Brothers:

$$CFTL_{i,t+26} = \alpha + \lambda \cdot (\ln(ILLIQ))_{i,t-35} + \Gamma' \cdot Z_{i,t-35} + e_{i,t+26}$$

There are 5 regressions in columns 1 through 5. The variables of each regression are described in the very left column. See Table 1 for the detailed definitions of the variables.

CFTLs are measured in percentage form. The sample consists of 769 stocks. t-statistics are inside the parentheses below the regression coefficients. 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)
ln(ILLIQ)	**-1.82*** (-3.30)**	**-1.65*** (-2.86)**	**-1.75*** (-3.06)**	**-2.41*** (-3.87)**	**-1.62** (-2.54)**
illiquidity beta	-	15.64 (1.09)	10.83 (0.76)	28.39* (1.90)	-14.48 (-0.01)
% of inst/nals	-	-	-14.74*** (-3.71)	-16.31*** (-3.97)	-23.96*** (-4.96)
market beta	-	-	-	19.06*** (4.10)	20.41*** (4.48)
ln(STF beta)	-	-	-	1.42 (0.65)	-7.58** (-2.50)
ln(mean turnover)	-	-	-	-	-10.95*** (-4.20)
ln(market-to-book)	-	-	-	-	0.57 (0.33)
momentum	-	-	-	-	-14.38** (-3.55)
debt-to-assets	-	-	-	-	0.02 (0.42)
RoA	-	-	-	-	0.54*** (2.31)
RoE	-	-	-	-	0.01 (0.27)
Adj-R^2 (%)	0.18%	0.10%	0.50%	3.19%	3.62%

10. Decomposition of ILLIQ to Size, Volatility and Turnover

10.1 Decomposition of ILLIQ and cross-sectional correlations

The ILLIQ measure is by construction correlated with the size of a stock, the volatility of the stock and its turnover.[81] These three variables define the level of the ILLIQ measure. Roughly, ILLIQ is the ratio of the absolute value of the return to the dollar volume $= \frac{|return|}{\$\,volume}$.[82] The numerator of the ratio is a volatility measure. Consider that the standard deviation of the stock returns (a measure of volatility) for a period is estimated as the square root of the variance, a number very close to the mean of the absolute values of the returns, for the estimation period. The denominator can be decomposed to the product of the size with the turnover of a stock:

$$\$volume = price \cdot (\#\ of\ shares\ traded) = \frac{size}{(total\ \#\ of\ shares)} \cdot (\#\ of\ shares\ traded) = size \cdot turnover \quad (15).$$

This is a direct algebraic connection of the denominator of ILLIQ, with the size and the turnover.

Amihud (2002) directly connects ILLIQ to the size of a stock (with the size we mean the market capitalization of a stock). We avoid including size and volatility as control variables to our main specifications of the cross-sectional regressions of CAR on ILLIQ for two reasons. The first reason is about the intuition of our analysis. We accept that a stock with low price change for a specific amount of trade is liquid. Thus, we can proxy illiquidity either by directly using the ILLIQ or by using its constituents. We prefer to use the ILLIQ measure as it includes both the effect of volatility and size. In addition, it is a variable widely used by the relevant literature and close to the sense of liquidity as it is perceived by the market participants. The second reason we do not include size and volatility as control variables to the cross-sectional regressions, is to avoid multicollinearity.

On the contrary we choose to include turnover as a control variable to the cross-sectional regressions, because we find that it has low correlation with ILLIQ and we want to

[81] Florackis et al. (2011) and Brennan et al. (2013) explain analytically the inherent relation of ILLIQ with the size of the stock.
[82] See also formula (9) about the estimation of ILLIQ

control for the level of trading activity, due to the use of the STF factor for the determination of the returns, before and during the crisis. Besides, there is evidence that turnover is a significant determinant of stock returns, over and above its multiplicative effect related to the cost per trade (Florackis et al (2011)).

We examine the cross-sectional correlations of the logarithm of ILLIQ with the logarithm of size, the logarithm of volatility and the logarithm of the share turnover (Table 5). The correlations of ln(ILLIQ) with ln(size), ln(volatility) and ln(turnover), are -0.98, 0.45 and -0.12, respectively. Big firms are more liquid, having lower $ILLIQ$ (negative sign of correlation). The correlation between the two variables is very strong and indicative for the significance of size to the illiquidity of a stock (both mechanically (through the construction of ILLIQ) and conceptually (taking into account the cross-sectional correlation of ILLIQ with other illiquidity measures, constructed with other methodologies). Volatile stocks are less liquid, as the positive sign between ILLIQ and volatility indicates. This is a well established stylized fact in the literature of market microstructure and market liquidity. Finally, the correlation between ILLIQ and turnover is lower and with negative sign. The negative sign is an expected result, as stocks with higher trading activity are easily traded and thus more liquid, while at the same time more liquid stocks are the target of investors with short horizons and higher turnover (Amihud and Mendelson (1986), Cella et al. (2013)).

The low correlation between ILLIQ and turnover is evidence that the STF and its coefficient (that we take from the equation (4)), capture in large extend different aspects of illiquidity. We are thus confident that the measurement of the abnormal trading activity during the crisis is not affected mechanically by a strong relation between liquidity and turnover. Besides, it is well established in the literature that the trading volume (and the turnover) does not necessarily measure liquidity.[83]

10.2 Cross-sectional regressions with size and volatility

Size and volatility are important determinants of the illiquidity of a stock. Each of the two variables is characterized by some special features that could give rise to different predictions about the relation between stock illiquidity and stock returns, during a large market decline.

[83] Although volume and liquidity are positively correlated.

A large size of a firm is translated into a large pool of investors that hold shares of it[84]. This in turn leads to fewer asymmetric information problems, lower search costs and - from an individual investor perspective - the opportunity to trade at a low cost. Moreover, for a given magnitude of transaction, a bigger size means lower proportion of the whole capitalization and lower price impact. Consequently, size is a natural candidate, as a stock characteristic, to predict flight-from-liquidity effect, since investors may prefer to sell big stocks to absorb funding liquidity. It is also expected that size predict flight-to-liquidity, as it is a main proxy of the liquidity of a stock.

On the contrary, volatility is a variable strongly connected with flight-to-liquidity (which in turns is connected with the flight-to-quality prediction, Vayanos (2004)). Increased stock price volatility leads to increased inventory risk, and as a result liquidity providers demand larger compensation to supply liquidity for a stock. In addition, fund managers, who are subject to internal risk controls, recognize that a stock with higher volatility may suffer a significant adverse price move that would force them to liquidate and write down losses. Furthermore, investors take into account that market illiquidity is persistent and prefer to hold more liquid stocks, to reduce the price impact of their transactions, in case they need to sell them in the future. Thus, we expect that volatility predicts flight-to-liquidity.

We test our hypothesis about the role of the determinants of ILLIQ, namely the size and the volatility, by running cross-sectional regressions of CAR, CFFL and CFTL on them. We report the results of the cross-sectional regressions on Table 9. The results strongly support our hypothesis about the role of size and volatility on the stock returns during the crisis. The two constituents of ILLIQ have different predictive patterns over the CAR and its two parts.

Size predicts flight-from-liquidity (negative sign in the second column, which means that bigger stocks lose more) but not flight-to-liquidity (insignificant coefficient in the third column, which means that size does not matter). Overall size has limited predictive ability over cumulative abnormal returns during the crisis, corroborating with a flight-from-liquidity phenomenon (negative sign and slightly significant coefficient in the first column).

On the contrary, volatility predicts flight-to-liquidity, as it enters the regression of the first column with a big, negative and statistically significant coefficient (more volatile stocks lose more). As expected, volatility enters into the regression of the CFTL (third column) with negative coefficient. This part of CAR is related with the flight-to-liquidity. Finally, the

[84] Or follow regularly the stock.

coefficient of the volatility in the regression of the CFFL (the part of CAR related with the abnormal selling) is again negative but statistically insignificant. That means that volatility does not lead to abnormal selling and if so, investors sell the more volatile stocks (again flight-to-liquidity). The insignificant coefficient of volatility on the second column combined with its statistical significance on the third column indicates that the effect of flight-to-liquidity is a price-risk effect, and does not affect the volume of trade. On the contrary, flight-from-liquidity is a volume effect, with size appearing with statistical significant coefficient on the regression of the second column.

Table 9: Cumulative abnormal stock returns (CAR), CFFL and CFTL, after Lehman's collapse and pre-crisis size, volatility and turnover (the parts of ILLIQ).

Cross-sectional Robust OLS regressions of the cumulative abnormal stock return (CAR) of stock i 19 weeks after the collapse of Lehman Brothers, $CAR_{i,t+26}$, on the natural logarithm of ILLIQ of stock i, $log(ILLIQ)_{i,t-35}$, as it is observed 15 weeks before the collapse of Lehman Brothers, and on other lagged control variables for stock i, $Z_{i,t-35}$, which are also observed 15 weeks before the collapse of Lehman Brothers:

$$CAR_{i,t+26} = \alpha + \lambda_{size} \cdot (\ln\,(size))_{i,t-35} + \lambda_{volatility} \cdot (\ln\,(volatility))_{i,t-35} + \Gamma' \cdot Z_{i,t-35} + e_{i,t+26}$$
$$CFFL_{i,t+26} = \alpha + \lambda_{size} \cdot (\ln\,(size))_{i,t-35} + \lambda_{volatility} \cdot (\ln\,(volatility))_{i,t-35} + \Gamma' \cdot Z_{i,t-35} + e_{i,t+26}$$
$$CFTL_{i,t+26} = \alpha + \lambda_{size} \cdot (\ln\,(size))_{i,t-35} + \lambda_{volatility} \cdot (\ln\,(volatility))_{i,t-35} + \Gamma' \cdot Z_{i,t-35} + e_{i,t+26}$$

There are 5 regressions in columns 1 through 5. The variables of each regression are described in the very left column. See Table 1 for the detailed definitions of the variables.

CARs are measured in percentage form. The sample consists of 769 stocks. t-statistics are inside the parentheses below the regression coefficients. 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 %.

	(CAR)	(CFFL)	(CFTL)
ln(size)	**-1.63***	**-1.73*****	**0.00**
	(-1.81)	**(4.07)**	**(~ 0.00)**
ln(volatility)	**-22.34*****	**-1.65**	**-17.77*****
	(-3.63)	**(-0.55)**	**(-3.15)**
ln(turnover)	-5.54	-1.29	-3.25
	(-1.42)	(-0.67)	(-0.90)
market beta	24.01***	0.88	25.15***
	(4.25)	(0.33)	(4.99)
illiquidity beta (*100)	-33.34*	-9.34	-5.78
	(-1.86)	(-1.10)	(-0.36)
% of inst/nals	-15.13***	0.93	-21.55***
	(-3.83)	(0.37)	(-4.52)
ln(STF beta)	-1.25	2.58	-3.18
	(-0.33)	(1.37)	(-0.90)
ln(market-to-book)	1.56	0.04	2.44
	(0.80)	(0.05)	(1.40)
momentum	-14.03***	4.13*	-15.97***
	(-3.03)	(1.89)	(-3.86)
debt-to-assets	-0.00	-0.06*	0.00
	(-0.06)	(-1.92)	(0.03)
RoA	0.85***	0.11	0.43*
	(3.21)	(0.91)	(1.85)
RoE	-0.08	-0.00	0.01
	(-1.31)	(-0.18)	(0.30)
Adj-R^2 (%)	8.57%	8.21%	9.15%

11. Robustness check

We repeat our main analysis using the percentage bid-ask spread as the illiquidity independent variable. Spread exhibits high cross-sectional correlation with ILLIQ, but it has not a price impact nature thus it is even more distant from the notion of the STF beta. The econometric specifications that we use are the same regarding all the control variables; hence the results are directly compared to that of the Tables 6, 7 and 8.

Table 10 reports the coefficients of the regression of CAR on ln(spread) and the control variables. The coefficient of ln(spread) is negative and insignificant in all cases, except from the specification of column (4), in which is significant. The significance of the coefficient in column (4) comes from the very high coefficient of ln(spread) in the CFTL part (column (4), Table 12). It seems that the lack of evidence about flight-to-liquidity in CAR does not stem from the use of ILLIQ as illiquidity proxy. The coefficients of the control variables are similar to that of Table 6 (basic results). Adjusted R-squares are also similar.

Table 11 reports the coefficients of the regression of CFFL on ln(spread) and the control variables. The coefficient of ln(spread) is positive and statistically significant in all specifications. It seems that the choice of illiquidity proxy does not affect the results regarding the flight-from-liquidity phenomenon. Again in this case, the coefficients for the remaining variables are similar to those of the basic analysis (Table 7). Finally, Table 12 reports the coefficients of the regression of CFTL on ln(spread) and the control variables. The coefficient of ln(spread) is negative and statistically significant in all specifications. The coefficients of the control variables remain in similar levels with those of Table 8.

Table 10: Cumulative abnormal stock returns (CAR) after Lehman's collapse and pre-crisis bid-ask spread

Cross-sectional Robust OLS regressions of the cumulative abnormal stock return (CAR) of stock i 19 weeks after the collapse of Lehman Brothers, $CAR_{i,t+26}$, on the natural logarithm of ILLIQ of stock i, $ln(spread)_{i,t-35}$, as it is observed 15 weeks before the collapse of Lehman Brothers, and on other lagged control variables for stock i, $Z_{i,t-35}$, which are also observed 15 weeks before the collapse of Lehman Brothers:

$$CAR_{i,t+26} = \alpha + \lambda \cdot (\ln{(spread)})_{i,t-35} + \Gamma' \cdot Z_{i,t-35} + e_{i,t+26}$$

There are 5 regressions in columns 1 through 5. The variables of each regression are described in the very left column. See Table 1 for the detailed definitions of the variables.
CARs are measured in percentage form. The sample consists of 769 stocks. t-statistics are inside the parentheses below the regression coefficients. 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)
ln(spread)	**-2.25** **(-1.21)**	**-2.43** **(-1.23)**	**-3.22** **(-1.62)**	**-5.82***** **(-2.72)**	**-3.45** **(-1.56)**
illiquidity beta (*100)	-	-5.45 (-0.35)	-11.01 (-0.70)	-2.82 (-0.17)	-21.06 (-0.87)
% of inst/nals	-	-	-12.13*** (-2.76)	-12.04*** (-2.64)	-15.66*** (-2.91)
market beta	-	-	-	13.95*** (2.69)	18.51*** (3.44)
ln(STF beta)	-	-	-	-4.40* (-1.90)	-8.65*** (-2.60)
ln(mean turnover)	-	-	-	-	-15.98*** (-5.47)
ln(market-to-book)	-	-	-	-	-0.48 (-0.25)
momentum	-	-	-	-	-10.98** (-2.41)
debt-to-assets	-	-	-	-	0.02 (0.28)
RoA	-	-	-	-	0.94*** (3.52)
RoE	-	-	-	-	-0.10 (-1.51)
Adj-R^2 (%)	0.03%	0.05%	1.20%	2.71%	7.97%

Table 11: Cumulative abnormal FFL part of returns (CFFL) after Lehman's collapse and pre-crisis bid-ask spread

Cross-sectional Robust OLS regressions of the cumulative FFL part of abnormal stock return (CFFL) of stock i 19 weeks after the collapse of Lehman Brothers, $CFFL_{i,t+26}$, on the natural logarithm of ILLIQ of stock i, *log(ILLIQ)*$_{i,t-35}$, as it is observed 15 weeks before the collapse of Lehman Brothers, and on other lagged control variables for stock i, $Z_{i,t-35}$, which are also observed 15 weeks before the collapse of Lehman Brothers:

$$CFFL_{i,t+26} = \alpha + \lambda \cdot (\ln\,(spread))_{i,t-35} + \Gamma' \cdot Z_{i,t-35} + e_{i,t+26}$$

There are 5 regressions in columns 1 through 5. The variables of each regression are described in the very left column. See Table 1 for the detailed definitions of the variables.

CFFLs are measured in percentage form. The sample consists of 769 stocks. t-statistics are inside the parentheses below the regression coefficients. 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)
ln(spread)	**4.06*** (4.65)**	**3.52*** (3.80)**	**3.57*** (3.82)**	**3.16*** (3.14)**	**3.55*** (3.37)**
illiquidity beta	-	-13.18* (-1.78)	-12.81* (-1.72)	-17.03** (-2.19)	-11.54 (-1.39)
% of inst/nals	-	-	0.85 (1.41)	2.11 (0.99)	1.62 (0.64)
market beta	-	-	-	-1.92 (-0.79)	0.67 (0.26)
ln(STF beta)	-	-	-	2.74** (2.48)	0.71 (0.44)
ln(mean turnover)	-	-	-	-	-2.47* (-1.76)
ln(market-to-book)	-	-	-	-	-0.29 (-0.33)
momentum	-	-	-	-	4.75** (2.21)
debt-to-assets	-	-	-	-	-0.07** (-2.10)
RoA	-	-	-	-	0.12 (1.01)
RoE	-	-	-	-	-0.00 (-0.24)
Adj-R^2 (%)	0.65%	0.64%	0.51%	2.63%	6.06%

Table 12: Cumulative abnormal FTL part of returns (CFTL) after Lehman's collapse and pre-crisis bid-ask spread

Cross-sectional Robust OLS regressions of the cumulative FTL part of abnormal stock return (CFTL) of stock i 19 weeks after the collapse of Lehman Brothers, $CFTL_{i,t+26}$, on the natural logarithm of ILLIQ of stock i, *log(ILLIQ)*$_{i,t-35}$, as it is observed 15 weeks before the collapse of Lehman Brothers, and on other lagged control variables for stock i, $Z_{i,q}$, which are also observed 15 weeks before the collapse of Lehman Brothers:

$$CFTL_{i,t+26} = \alpha + \lambda \cdot (\ln{(spread)})_{i,t-35} + \Gamma' \cdot Z_{i,t-35} + e_{i,t+26}$$

There are 5 regressions in columns 1 through 5. The variables of each regression are described in the very left column. See Table 1 for the detailed definitions of the variables.

CFTLs are measured in percentage form. The sample consists of 769 stocks. t-statistics are inside the parentheses below the regression coefficients. 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)
ln(spread)	**-6.00*** (-3.52)**	**-5.49*** (-3.05)**	**-6.17*** (-3.44)**	**-8.55*** (-4.42)**	**-5.19*** (-2.62)**
illiquidity beta	-	14.58 (1.02)	8.83 (0.62)	26.84* (1.80)	-38.19 (-0.02)
% of inst/nals	-	-	-15.41*** (-3.87)	-17.33*** (-4.22)	-24.39*** (-5.12)
market beta	-	-	-	20.38*** (4.37)	20.65*** (4.32)
ln(STF beta)	-	-	-	1.34 (0.63)	-7.19** (-2.34)
ln(mean turnover)	-	-	-	-	-10.19*** (-3.87)
ln(market-to-book)	-	-	-	-	0.79 (0.46)
momentum	-	-	-	-	-14.64*** (-3.63)
debt-to-assets	-	-	-	-	0.03 (0.56)
RoA	-	-	-	-	0.55** (2.33)
RoE	-	-	-	-	0.01 (0.22)
Adj-R^2 (%)	1.12%	0.99%	1.58%	4.94%	5.33%

12. Conclusion

We study the relation between stock illiquidity and the cross-section of stock returns, during a large market decline. According to the related literature, liquid stocks perform better compared to illiquid stocks during a crisis (flight-to-liquidity phenomenon). The reason is the shift of the investors' preferences towards liquidity to minimize (a) the possibility of large adverse price movements in the continuation of the crisis and (b) the price impact, in case they need to liquidate.

Nevertheless, a number of empirical papers which use data from the financial crisis of 2007-2009 show that hedge funds and mutual funds prefer to sell not only illiquid stocks, but also liquid stocks to obtain funding liquidity at the minimum possible price discount. These initial observations motivated us to study the cross-section of the stock returns during the crisis, in respect to the illiquidity levels of the stocks.

Our empirical approach introduced a signed turnover factor that captured the trading pressure for a stock for each week, and was then included in a standard four-factor model of stock returns as a predictor variable. The new factor gave us the opportunity to identify two components, namely normal and abnormal trading. We were then able to measure the abnormal returns during a crisis (as the difference of the realized returns minus the prediction of our five factor model, where as input of the fifth factor, we used only its normal part). Furthermore, we could split the abnormal returns during a crisis into one part attributable to abnormal trading (selling in our case) and into a second part attributable to additional effects. The former part is related to a flight-from-liquidity phenomenon, a new concept observed in this study, which states that investors sell liquid stocks to absorb funding liquidity during a crisis. The latter part (additional effects) is related to the classical flight-to-liquidity prediction. This motivated us to test the unbiased cumulative abnormal returns during the crisis and examine the simultaneous existence of two opposite phenomena, both stemming from the nature of stock illiquidity.

To check if illiquidity is a cross-sectional predictor of stock returns during a crisis, we ran cross-sectional regressions of the cumulative abnormal returns and their two parts, on the illiquidity measure ILLIQ (Amihud, 2002), illiquidity beta (that measures the illiquidity risk of a stock) and other relevant control variables. Our results indicated that it is not the case. The reason was that, except from the flight to liquidity effect, the flight from liquidity phenomenon is also prevalent. The correct measurement of the abnormal returns allowed us to identify both effects. Liquid stocks performed better than illiquid ones in terms of pricing due to their individual level of liquidity but performed worse than illiquid stocks in terms of the price impact of the abnormal selling activity for funding reasons.

We then tested the ability of illiquidity risk (measured by the illiquidity beta) to make cross-sectional predictions of the stock returns during a crisis. We showed that illiquidity beta fails to predict stock returns, because it is also related with the existence of the two opposite phenomena and because it lost its statistical significance when we add market beta as a regressor.

In addition, we decomposed ILLIQ into its component variables, size, volatility and turnover. Size and volatility are well known liquidity proxies. Volatility is connected with the flight to liquidity prediction, but size is connected with both the liquidity effects. Our results from the cross-sectional regressions support the different patterns of size and volatility. Size is the main driver of the flight-from-liquidity effect. Investors prefer to absorb funding liquidity by selling large stocks to minimize their adverse price impact from the transactions.

Our results call for further theoretical and empirical investigations regarding the role of stock illiquidity on stock returns during a crisis. They are also useful for risk management purposes, especially for investors and funds that need to take into account the liquidity of their investment. Finally, our results contribute to the research agenda of the measurement of the systemic risk, providing a systemic empirical analysis of the role of stock liquidity during a crisis.

Appendix A: Definitions of the variables of the paper

Table A.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
Signed Turnover Factor ($STF_{i,t}$)	The STF of stock i during week t is measured as the sum of the number of shares of stock i that are traded at each day d of week t, after taking the sign of the return of stock i, at the respective day d, and divided by the total number of shares of stock i during week t: $STF_{i,t} = \frac{\sum_{d=1}^{5} sign(return)_{i,d} * (\#\ of\ shares\ traded)_{i,d}}{(total\ \#\ of\ shares)_{i,t}}$	We take the trading volume, the total number of shares and the stock returns from Bloomberg. (Bloomberg Datatypes: PX_VOLUME, EQY_SH_OUT and DAY_TO_DAY_TOT_RETURN_GROSS_DVDS, respectively)	335 weeks, 769 firms
Signed Normal STF ($norSTF_{i,t}$)	The signed normal STF for stock i at the end of week t, is the normal level of STF for stock i for week t, after taking the sign of the STF of the same week. The normal level of STF for stock for week t is the mean of the absolute value of STF of the weeks t-52 to t, after adjusting for a time trend: $(normal\ level\ of\ STF)_{i,t} = a_{i,t}^{STF} + b_{i,t}^{week} \cdot t$ The mean absolute value of STF ($a_{i,t}^{STF}$) and the coefficient of the time trend ($b_{i,t}^{week}$) are retrieved from rolling time series regressions (with a window from t-52 to t-1) of the absolute value of STF for stock i on the number of the weeks: $(\lvert STF \rvert)_{i,t} = a_i^{STF} + b_i^{week} \cdot t + e_{i,t}^{STF}$	We take the trading volume, the total number of shares and the stock returns from Bloomberg. (Bloomberg Datatypes: PX_VOLUME, EQY_SH_OUT and DAY_TO_DAY_TOT_RETURN_GROSS_DVDS, respectively)	283 weeks, 769 firms
Abnormal Level of STF ($abnSTF_{i,t}$)	The abnormal level of STF for stock i at the end of week t, is the difference of the realized STF for week t minus the Signed Normal STF for the same week.	We take the trading volume, the total number of shares and the stock returns from Bloomberg. (Bloomberg Datatypes: PX_VOLUME, EQY_SH_OUT and DAY_TO_DAY_TOT_RETURN_GROSS_DVDS, respectively)	283 weeks, 769 firms
Excess market Return	The excess market return is the value-weight return of all CRSP stocks that are incorporated in the US and are listed on NYSE,	Rm-Rf directly from the site of Kenneth French:	335 weeks

Rm_{q+1}-Rf_q	AMEX or NASDAQ and have share code 10 or 11 minus the risk-free rate (Treasury bill rate) for the relevant period.	http://mba.tuck.dartmouth.edu/pages/faculty/ken.french/data_library.html#Research	
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	335 weeks
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	335 weeks
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).	We construct MOM factor with data from our sample.	335 weeks
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.	335 weeks
market beta / SMB beta / HML beta / MOM beta / STF beta	Betas from rolling time-series regressions (with a 52-week window) of the weekly excess stock returns on the following five factors: Excess market return (Rm-Rf), SMB (Small-minus-Big), HML (High-minus-Low) and MOM (winner-minus-losers), STF: $r_{i,t} - r_t^f = a + b_i^m(R_m - r^f)_t + b_i^{smb}(SMB)_t + b_i^{hml}(HML)_t + b_i^{mom}(MOM)_t + b_i^{STF}(STF)_{i,t} + e_{i,t}$. We measure the weekly excess stock returns by subtracting from	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)	283 weeks, 769 firms, of each of the betas

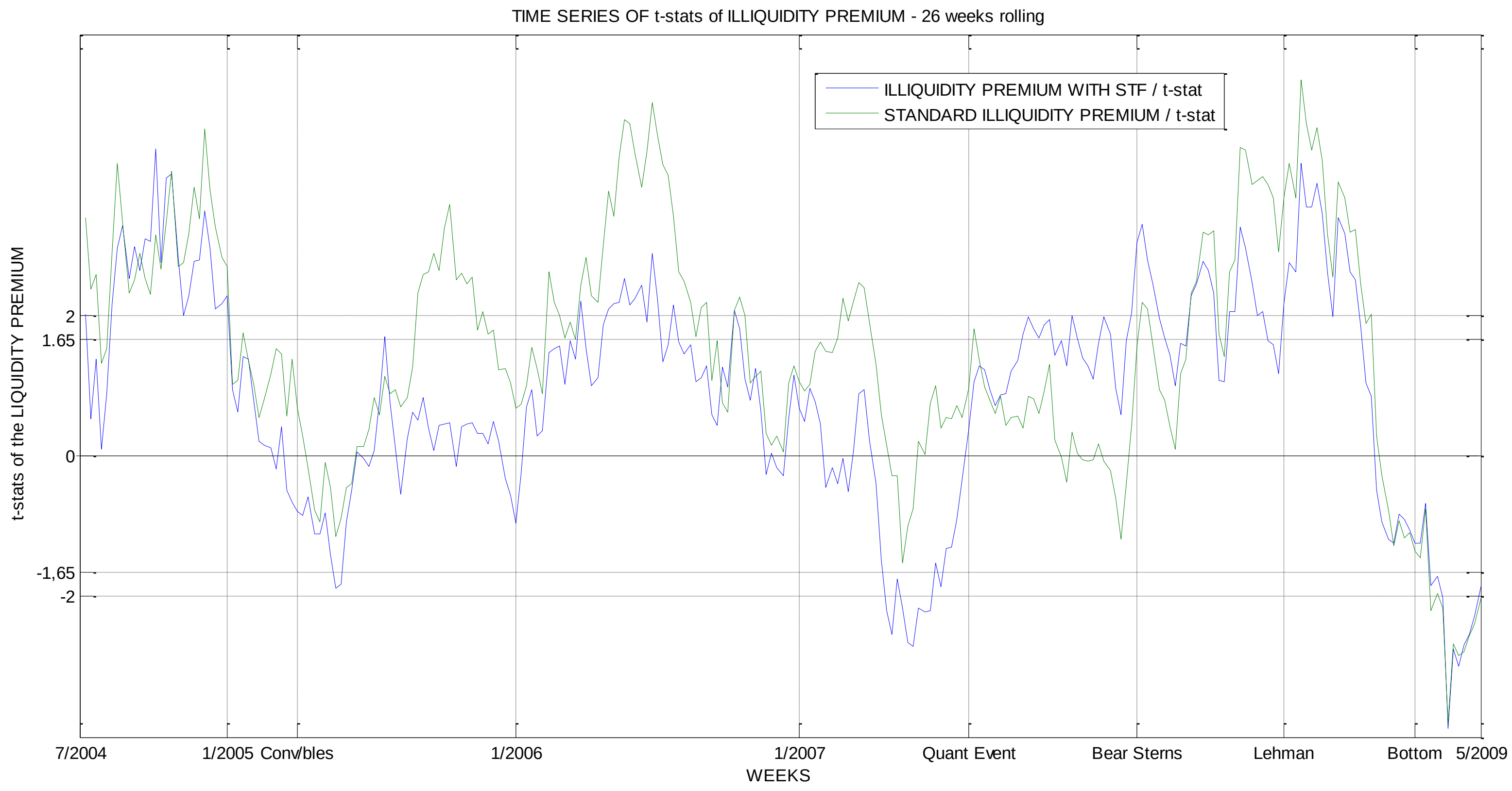
	the weekly stock price changes the risk-free rate. We use the natural logarithm of the b_i^{STF} to our econometric specifications.		
Cumulative Abnormal Returns ($CAR_{i,t}$)	The Cumulative Abnormal Returns (CAR) of stock i at week t, is the sum of the abnormal returns (AR) of the stock from the end of the week t-26 until the end of the week t. The abnormal return of stock i for week t ($AR_{i,t}$), is the difference between the realized return of the week minus the prediction of a 5-factor model for the same week: $AR_{i,t} = r_{i,t} - \left[\widehat{b_i^m} \cdot (R_m - r_f)_t + \widehat{b_i^{SMB}} \cdot SMB_t + \widehat{b_i^{HML}} \cdot HML_t + \widehat{b_i^{MOM}} \cdot MOM_t + \widehat{b_i^{STF}} \cdot (norSTF)_{i,t}\right]$	Bloomberg. (Bloomberg Datatype: DAY_TO_DAY_TOT_RETURN_GROSS_DVDS)	257 weeks, 769 firms
Cumulative Flight From Liquidity part of Abnormal Returns ($CFFL_{i,t}$)	The Cumulative Flight From Liquidity part of Abnormal Returns (CFFL) of stock i at week t, is the sum of the Flight From Liquidity part of abnormal returns (FFL) of the stock from the end of the week t-26 until the end of the week t. The Flight From Liquidity part of abnormal return of stock i for week t ($FFL_{i,t}$), is the difference between the realized STF impact to the return of the week minus the predicted impact for the same week: $FFLr_{i,t} = \hat{b}_{5,i} \cdot (abnorSTF)_{i,t}$	We take the trading volume, the total number of shares and the stock returns from Bloomberg. (Bloomberg Datatypes: PX_VOLUME, EQY_SH_OUT and DAY_TO_DAY_TOT_RETURN_GROSS_DVDS, respectively)	257 weeks, 769 firms
Cumulative Flight To Liquidity part of Abnormal Returns ($CFTL_{i,t}$)	The Cumulative Flight From Liquidity part of Abnormal Returns (CFFL) of stock i at week t, is the sum of the Flight From Liquidity part of abnormal returns (FFL) of the stock from the end of the week t-26 until the end of the week t. The Flight From Liquidity part of abnormal return of stock i for week t ($FFL_{i,t}$), is the difference between the abnormal return ($AR_{i,t}$) of the week and the Flight From Liquidity part of abnormal return ($FFL_{i,t}$) of the same week: $FTLr_{i,t} = AR_{i,t} - \hat{b}_{5,i} \cdot (abnorSTF)_{i,t}$	We take the trading volume, the total number of shares and the stock returns from Bloomberg. (Bloomberg Datatypes: PX_VOLUME, EQY_SH_OUT and DAY_TO_DAY_TOT_RETURN_GROSS_DVDS, respectively)	257 weeks, 769 firms
ILLIQ (Amihud,2002) $\ln(ILLIQ)_{i,t}$	The natural logarithm of the ILLIQ measure. ILLIQ of stock i for week t is the average of the daily ratios of the absolute level of the stock price change to the dollar volume, multiplied by a scaling	Stock prices from Bloomberg. (Bloomberg Datatype: PX_LAST) Share volumes from Bloomberg.	335 weeks, 769 firms

	factor of 10^6: $ILLIQ_{i,t} = 1/D \cdot \sum_{d=1}^{D} \lvert r_{i,d} \rvert / \$volume_{i,d} \cdot 10^6$, where D is the total number of trading days of the period between t-52 to t.	(Bloomberg Datatype: PX_VOLUME)	
Bid-ask spread ln(spread)$_{i,t}$	The natural logarithm of the Bid-Ask Spread. Spread of stock i for week t is the average of the daily spread, as a percentage of stock price, of the period between t-52 to t.	Stock prices, bid prices and ask prices from Bloomberg.	335 weeks, 769 firms
Size ln(size)$_{i,t}$	The natural logarithm of market capitalization of stock i at the end of week t.	Bloomberg. (Bloomberg Datatype: CUR_MKT_CAP)	335 weeks, 769 firms
Volatility ln(volatility)$_{i,t}$	The natural logarithm of the standard deviation of the daily stock returns within the period starting at t-52 and ending at t.	Prices from Bloomberg. (Bloomberg Datatype: PX_LAST)	335 weeks, 769 firms
Share Turnover ln(turnover)$_{i,t}$	The natural logarithm of the share turnover of stock i for week t, is the average of the daily ratios of the number of shares traded to the total outstanding number of shares: $turnover_{i,t} = 1/D \cdot \sum_{d=1}^{D} volume_{i,d} / (total\ \#\ of\ shares)_{i,d}$, where D is the total number of trading days of the period between t-52 to t.	We take the trading volume and the total number of shares from Bloomberg. (Bloomberg Datatypes: PX_VOLUME and EQY_SH_OUT, respectively)	335 weeks, 769 firms
(illiquidity beta)$_{i,t}$	Illiquidity beta from rolling time-series regressions (with a 52-week 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: $r_{i,t} - r_t^f = a + b_i^{illiq}(innov - mILLIQ)_t + b_i^m (R_m - r^f)_t + e_{i,t}$. The $mILLIQ$ is the cross-sectional mean of the $ILLIQ$, for each week t. The innovations of $mILLIQ$ are the residuals of an AR(1) model: $(mILLIQ)_t = c + (mILLIQ)_{t-1} + (innov - mILLIQ)_t$. The illiquidity beta is multiplied by a scaling factor of 100.	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)	283 weeks, 769 firms
Level of Institutional Ownership ln(inst-perc)$_{i,q}$	The natural logarithm of the percentage level of institutional ownership for stock i at the end of quarter q.	Data from Thomson One, 13F Institutional Ownership Mode. We use data from the end of the second quarter of 2008.	1 quarter, 769 firms
Market-to-Book ln(mtb)$_{i,t}$	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	Market-to-Book ratios are directly provided by Bloomberg. (Bloomberg	335 weeks, 769 firms

	end of week t and book value is the accounting value of the stock i at the end of the previous year.	Datatype: MARKET_CAPITALIZATION_TO_BV)	
Price Momentum $mom_{i,t}$	The cumulative stock return measured over 11 months, from the end of month m-12 to the end of m-1, respective to each week: $mom_{i,t} = \frac{Price_{i,m-1} - Price_{i,m-12}}{Price_{i,m-12}}$	Prices from Bloomberg. (Bloomberg Datatype: PX_LAST)	257 weeks, 769 firms
Debt-to-Assets $\ln(dta)_{i,t}$	The natural logarithm of the ratio of total debt to total assets of stock i at the end of week t.	Debt-to-Assets ratios provided directly by Bloomberg. (Bloomberg Datatype: TOT_DEBT_TO_TOT_ASSET)	335 weeks, 769 firms
Return on Assets $RoA_{i,t}$	The ratio of earnings to total assets of stock i at the end of week t.	Return-on-Assets ratios provided directly by Bloomberg. (Bloomberg Datatype: ROA)	335 weeks, 769 firms
Return on Equity $RoE_{i,t}$	The ratio of earnings to shareholders' equity of stock i at the end of week t.	Return-on-Equity ratios provided directly by Bloomberg. (Bloomberg Datatype: ROE)	335 weeks, 769 firms

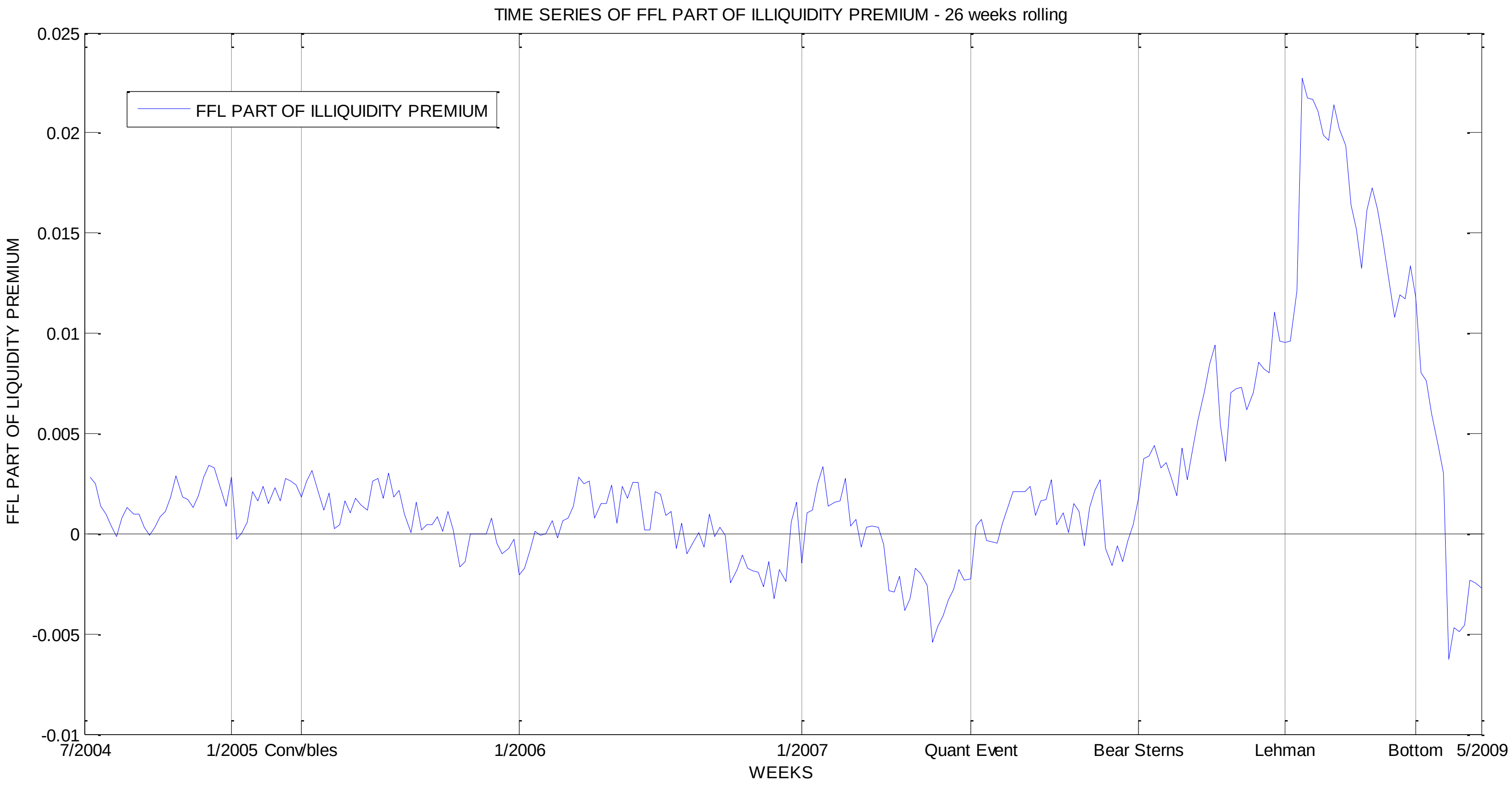
Appendix B: Supplementary figures from the whole period analysis

Figure 1.B: Level of statistical significance of illiquidity premium (γ coefficient): Classical approach Vs STF augmented approach.



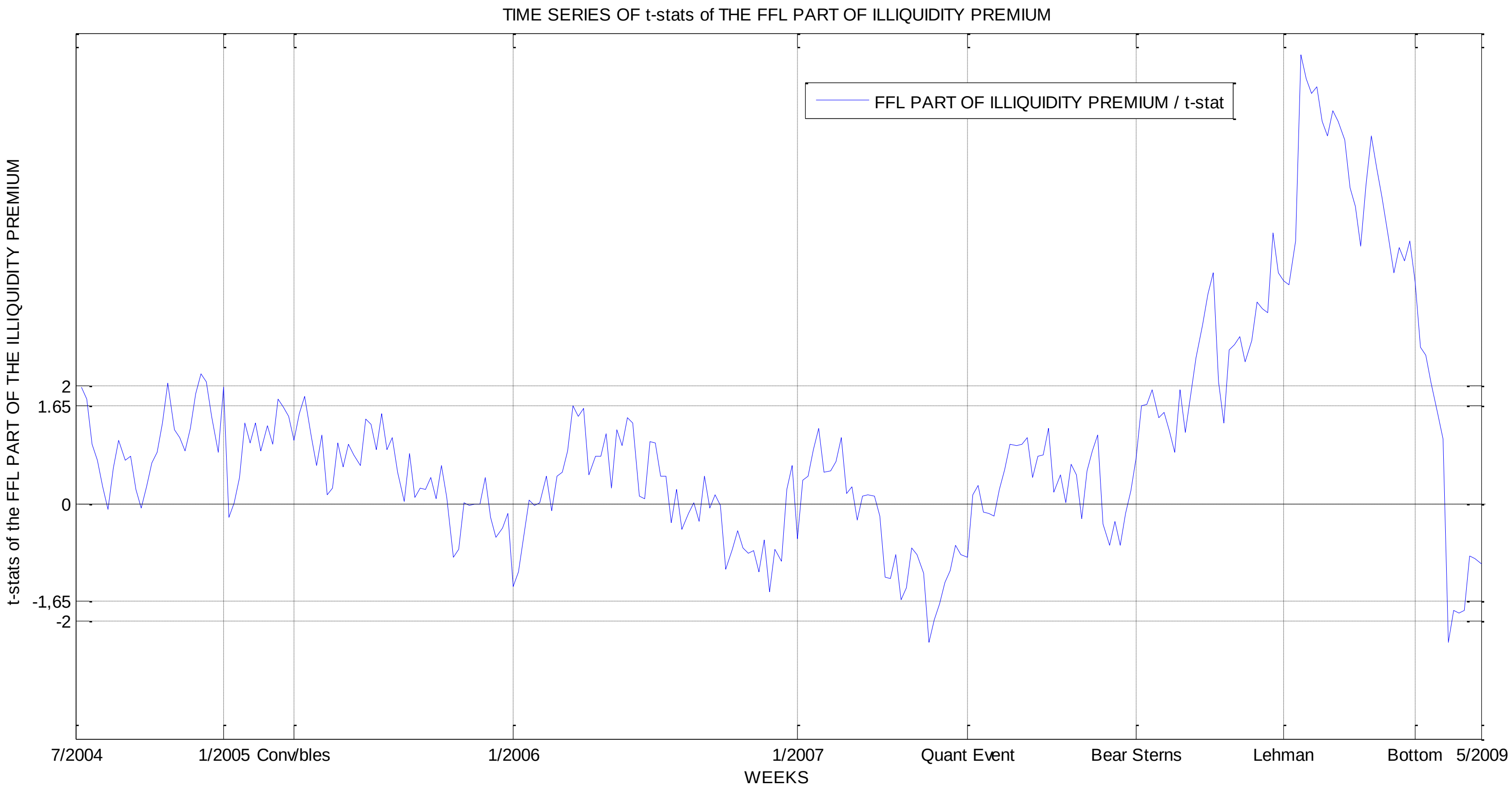
The figure illustrates the evolution of the t-statistics of the illiquidity premium which is estimated using both the classical approach (green line) and STF augmented approach (blue line), over the last 284 weeks of the sample (July 2004 to May 2009). Under the classical approach, the illiquidity premium is the coefficient (γ) of rolling cross-sectional regressions of the risk-adjusted returns (4-factor model) of each week on the ILLIQ of the 52 previous weeks. Under our approach, the illiquidity premium is the coefficient (γ) of rolling cross-sectional regressions of the risk and turnover adjusted returns (4-factor model plus STF) of each week on the ILLIQ of the 52 previous weeks.

Figure 3.B: The FFL part of the illiquidity premium (coefficient γ_{FFL}).



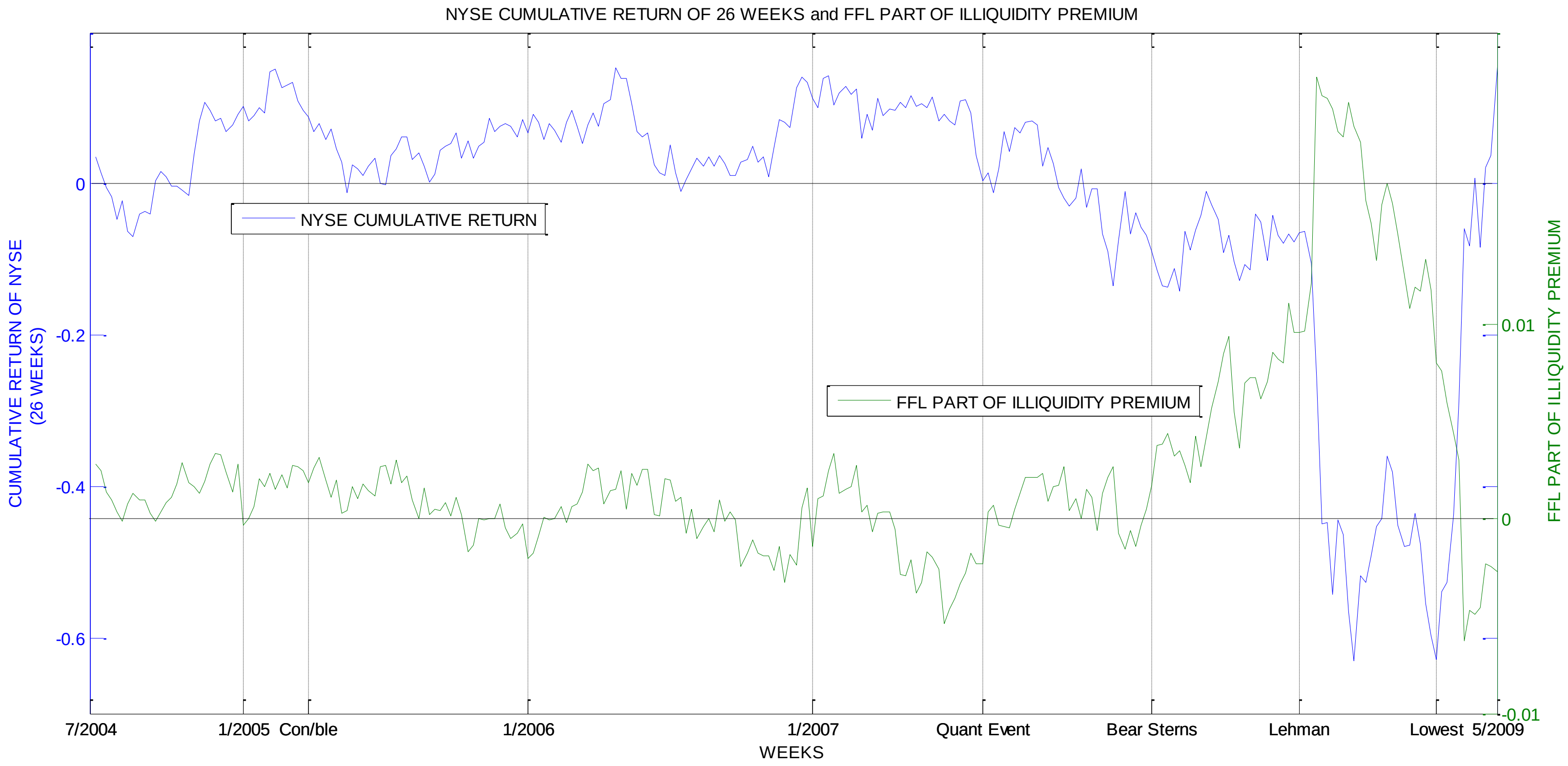
The figure illustrates the rolling FFL part of the illiquidity premium (coefficient γ_{FFL}), over the last 257 weeks of the sample (July 2004 to May 2009). The FFL part of the illiquidity premium is the coefficient (γ_{FFL}) of rolling cross-sectional regressions of the abnormal STF-related returns of each week on the ILLIQ of the 52 previous weeks. The abnormal STF-related returns are the product of the abnormal STF with the pre-estimated elasticity of stock returns on STF (b_{stf}).

Figure 3.C: Level of statistical significance of the FFL part of the illiquidity premium (coefficient γ_{FFL}).



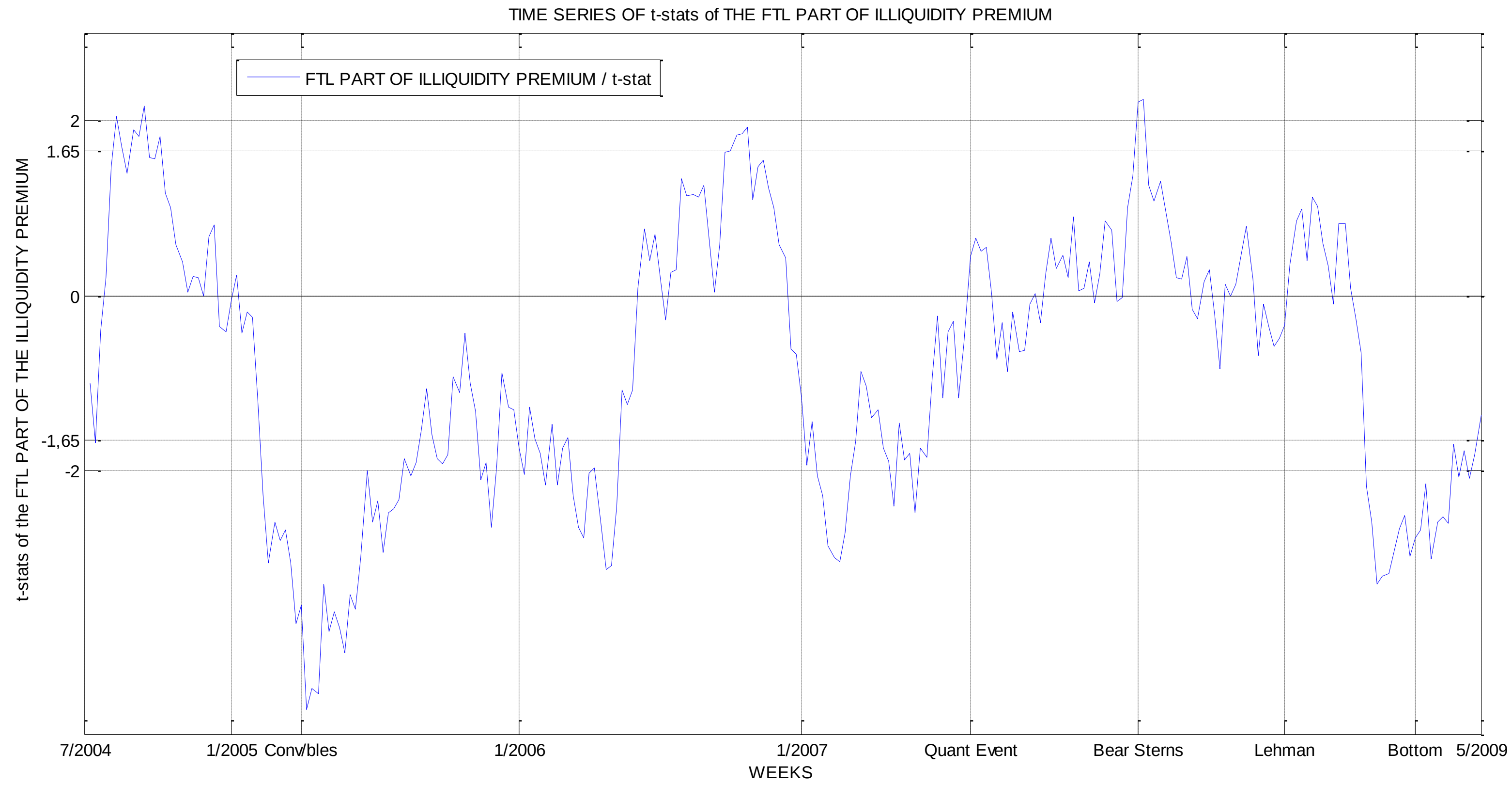
The figure illustrates the t-statistics of the rolling FFL part of the illiquidity premium (coefficient γ_{FFL}), over the last 257 weeks of the sample (July 2004 to May 2009). The FFL part of the illiquidity premium is the coefficient (γ_{FFL}) of rolling cross-sectional regressions of the abnormal STF-related returns of each week on the ILLIQ of the 52 previous weeks. The abnormal STF-related returns are the product of the abnormal STF with the pre-estimated elasticity of stock returns on STF (b_{stf}).

Figure 3D: The FFL part of the illiquidity premium (γ_{FFL} coefficient) (green line, right Y-axis) and the cumulative 26-week returns of NYSE composite index (blue line, left Y-axis).



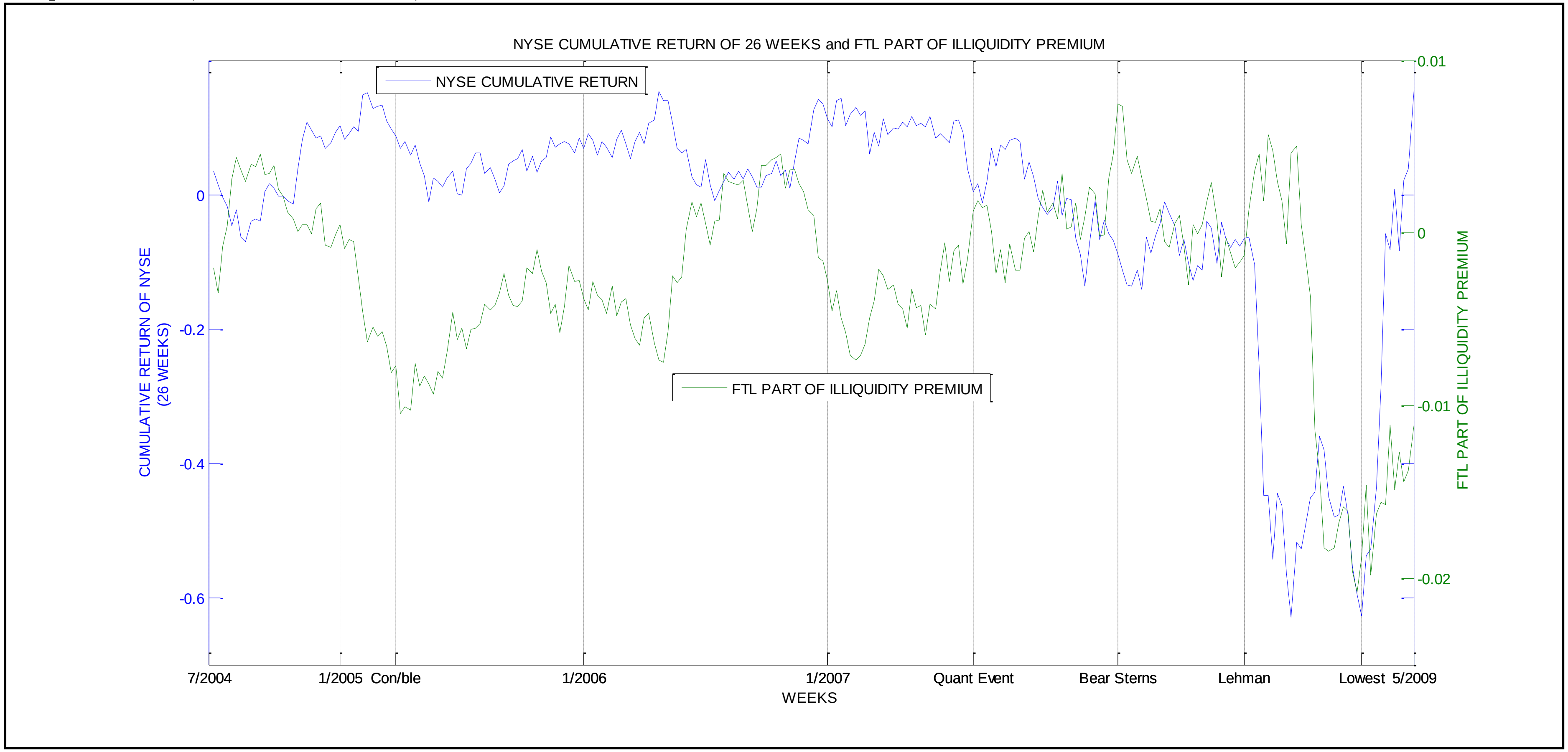
The figure illustrates the rolling FFL part of the illiquidity premium (coefficient γ_{FFL}) (green line, right Y-Axis) and the rolling cumulative 26-week return of NYSE composite index, over the last 257 weeks of the sample (July 2004 to May 2009). The FFL part of the illiquidity premium is the coefficient (γ_{FFL}) of rolling cross-sectional regressions of the abnormal STF-related returns of each week on the ILLIQ of the 52 previous weeks. The abnormal STF-related returns are the product of the abnormal STF with the pre-estimated elasticity of stock returns on STF (b_{stf}).

Figure 7.B: Level of statistical significance of the FTL part of the illiquidity premium (coefficient γ_{FTL}).



The figure illustrates the t-statistics of the rolling FTL part of the illiquidity premium (coefficient γ_{FTL}), over the last 257 weeks of the sample (July 2004 to May 2009). The FTL part of the illiquidity premium is the coefficient (γ_{FTL}) of rolling cross-sectional regressions of the risk and turnover adjusted returns (after the subtraction of the abnormal STF-related returns) of each week on the ILLIQ of the 52 previous weeks.

Figure 7.C: The FTL part of the illiquidity premium (γ_{FTL} coefficient) (green line, right Y-axis) and the cumulative 26-week returns of NYSE composite index (blue line, left Y-axis).



The figure illustrates FTL part of the illiquidity premium (coefficient γ_{FTL}) (green line, right Y-Axis) and the rolling cumulative 26-week return of NYSE composite index, over the last 257 weeks of the sample (July 2004 to May 2009). The FTL part of the illiquidity premium is the coefficient (γ_{FTL}) of rolling cross-sectional regressions of the risk and turnover adjusted returns (after the subtraction of the abnormal STF-related returns) of each week on the ILLIQ of the 52 previous weeks.

Appendix C: Supplementary tables from the main econometric analysis

Table 6b: Cumulative abnormal stock returns (CAR) after Lehman's collapse and pre-crisis ILLIQ

Cross-sectional Robust OLS regressions of the cumulative abnormal stock return (CAR) of stock i 26 weeks after the collapse of Lehman Brothers, $CAR_{i,t+26}$, on the natural logarithm of ILLIQ of stock i, *log(ILLIQ)*$_{i,t-15}$, as it is observed 15 weeks before the collapse of Lehman Brothers, and on other lagged control variables for stock i, $Z_{i,t-15}$, which are also observed 15 weeks before the collapse of Lehman Brothers:

$$CAR_{i,t+26} = \alpha + \lambda \cdot (\ln\,(ILLIQ))_{i,t-15} + \Gamma' \cdot Z_{i,t-15} + e_{i,t+26}$$

There are 5 regressions in columns 1 through 5. The variables of each regression are described in the very left column. See Table 1 for the detailed definitions of the variables.

CARs are measured in percentage form. The sample consists of 769 stocks. t-statistics are inside the parentheses below the regression coefficients. 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)
ln(ILLIQ)	**-1.23** **(-1.35)**	**-1.43** **(-1.64)**	**-1.51*** **(-1.73)**	**-2.03**** **(-2.14)**	**-2.02**** **(-2.03)**
illiquidity beta	-	-32.25 (-1.49)	-33.75 (-1.55)	0.32 (0.01)	-20.46 (-0.84)
% of inst/nals	-	-	-6.77 (-1.11)	-12.64** (-2.02)	-17.17*** (-2.31)
market beta	-	-	-	33.11*** (4.68)	40.07*** (5.39)
ln(STF beta)	-	-	-	-5.26 (-1.62)	-16.35*** (-3.57)
ln(mean turnover)	-	-	-	-	-14.33*** (-3.58)
ln(market-to-book)	-	-	-	-	-2.85 (-1.06)
momentum	-	-	-	-	-7.01 (-1.11)
debt-to-assets	-	-	-	-	-0.07 (-0.69)
RoA	-	-	-	-	1.29*** (3.47)
RoE	-	-	-	-	-0.16* (-1.69)
Adj-R^2 (%)	0.11%	0.99%	1.05%	4.15%	7.66%

Table 7b: Cumulative abnormal FFL part of returns (CFFL) after Lehman's collapse and pre-crisis ILLIQ

Cross-sectional Robust OLS regressions of the cumulative FFL part of abnormal stock return (CFFL) of stock i 26 weeks after the collapse of Lehman Brothers, $CAFFL_{i,26}$, on the natural logarithm of ILLIQ of stock i, $log(ILLIQ)_{i,t-15}$, as it is observed 15 weeks before the collapse of Lehman Brothers, and on other lagged control variables for stock i, $Z_{i,t-15}$, which are also observed 15 weeks before the collapse of Lehman Brothers:

$$CFFL_{i,t+26} = \alpha + \lambda \cdot (\ln{(ILLIQ)})_{i,t-15} + \Gamma' \cdot Z_{i,t-15} + e_{i,t+26}$$

There are 5 regressions in columns 1 through 5. The variables of each regression are described in the very left column. See Table 1 for the detailed definitions of the variables.

CFFLs are measured in percentage form. The sample consists of 769 stocks. t-statistics are inside the parentheses below the regression coefficients. 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)
ln(ILLIQ)	**1.17*** (3.79)**	**1.07*** (3.29)**	**1.08*** (3.30)**	**1.08*** (3.03)**	**1.06*** (2.80)**
illiquidity beta	-	-7.86 (-0.98)	-7.54 (-0.93)	-7.47 (-0.87)	-8.83 (-0.96)
% of inst/nals	-	-	0.65 (0.29)	0.63 (0.27)	2.24 (0.79)
market beta	-	-	-	0.02 (0.01)	0.92 (0.33)
ln(STF beta)	-	-	-	-0.04 (-0.04)	-2.33 (-1.30)
ln(mean turnover)	-	-	-	-	-2.54* (-1.64)
ln(market-to-book)	-	-	-	-	0.69 (0.68)
momentum	-	-	-	-	2.58 (1.07)
debt-to-assets	-	-	-	-	-0.09** (-2.33)
RoA	-	-	-	-	0.05 (0.35)
RoE	-	-	-	-	-0.02 (-0.69)
Adj-R^2 (%)	0.26%	0.15%	0.02%	1.55%	3.50%

Table 8b: Cumulative abnormal FTL part of returns (CFTL) after Lehman's collapse and pre-crisis ILLIQ

Cross-sectional Robust OLS regressions of the cumulative FTL part of abnormal stock return (CFTL) of stock i 26 weeks after the collapse of Lehman Brothers, $CAFTL_{i,26}$, on the natural logarithm of ILLIQ of stock i, *log(ILLIQ)*$_{i,t-15}$, as it is observed 15 weeks before the collapse of Lehman Brothers, and on other lagged control variables for stock i, $Z_{i,t-15}$, which are also observed 15 weeks before the collapse of Lehman Brothers:

$$CFTL_{i,t+26} = \alpha + \lambda \cdot (\ln\,(ILLIQ))_{i,t-15} + \Gamma' \cdot Z_{i,t-15} + e_{i,t+26}$$

There are 5 regressions in columns 1 through 5. The variables of each regression are described in the very left column. See Table 1 for the detailed definitions of the variables.

CFTLs are measured in percentage form. The sample consists of 769 stocks. t-statistics are inside the parentheses below the regression coefficients. 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)
ln(ILLIQ)	**-2.05*** (-2.73)**	**-2.49*** (-3.15)**	**-2.57*** (-3.25)**	**-3.22** (-3.75)**	**-2.70** (-3.08)**
illiquidity beta	-	-33.70* (-1.72)	-35.38* (-1.80)	-2.95 (-0.14)	-21.55 (-1.01)
% of inst/nals	-	-	-7.53 (-1.37)	-12.37** (-2.20)	-10.87 (-1.51)
market beta	-	-	-	32.54*** (5.09)	38.01*** (5.79)
ln(STF beta)	-	-	-	-3.51 (-1.19)	-11.89*** (-2.95)
ln(mean turnover)	-	-	-	-	-11.41*** (-3.22)
ln(market-to-book)	-	-	-	-	-4.84** (-1.96)
momentum	-	-	-	-	-12.39** (-2.22)
debt-to-assets	-	-	-	-	0.04 (0.53)
RoA	-	-	-	-	1.36*** (4.14)
RoE	-	-	-	-	0.01 (0.11)
Adj-R^2 (%)	0.63%	1.47%	1.58%	4.48%	7.07%

References

Acharya Viral V., Pedersen Lasse Heje, 2005, Asset pricing with liquidity risk, *Journal of Financial Economics* 77, 375-410.

Amihud, Yakov, 2002, Illiquidity and stock returns: Cross-section and time-series effects, *Journal of Financial Markets* 5, 31-56.

Amihud, Yakov, and Haim Mendelson, 1986, Asset pricing and the bid-ask spread, *Journal of Financial Economics* 17, 223-249.

Anshuman, Ravi and S. Viswanathan, 2005, Costly collateral and liquidity, *Working paper, Duke University*.

Attari, Mukarram, Antonio S. Mello, and Martin E. Ruckes, 2005, Arbitraging arbitrageurs, *The Journal of Finance* 60, 2471-2511.

Ben-David, Itzhak, Francesco Franzoni, and Rabih Moussawi, 2012, Hedge fund stock trading in the financial crisis of 2007–2009, *Review of Financial Studies* 25, 1-54.

Bernardo, Antonio E., and Ivo Welch, 2004, Liquidity and financial market runs, *The Quarterly Journal of Economics* 119, 135-158.

Blume, Marshall E., A. Craig Mackinlay, and Bruce Terker, 1989, Order imbalances and stock price movements on october 19 and 20, 1987, *The Journal of Finance* 44, 827-848.

Brennan, Michael, Sahn-Wook Huh, and Avanidhar Subrahmanyam, 2013, An analysis of the amihud illiquidity premium, *Review of Asset Pricing Studies* 3, 133-176.

Brennan Michael J., Chordia Tarun, Subrahmanyam Avanidhar, Tong Qing 2012, Sell-order liquidity and the cross-section of expected stock returns, *Journal of Financial Economics* 105, 523-541.

Brunnermeier, Markus K., 2009, Deciphering the liquidity and credit crunch 2007-2008, *Journal of Economic Perspectives* 23, 77-100.

Brunnermeier, Markus K., and Lasse Heje Pedersen, 2005, Predatory trading, *The Journal of Finance* 60, 1825-1863.

Brunnermeier, Markus K., and Lasse Heje Pedersen, 2009, Market liquidity and funding liquidity, *Review of Financial Studies* 22, 2201-2238.

Cella, Cristina, Andrew Ellul, and Mariassunta Giannetti, 2013, Investors' horizons and the amplification of market shocks, *Review of Financial Studies*.

Chordia, Tarun, Richard Roll, and Avanidhar Subrahmanyam, 2000, Commonality in liquidity, *Journal of Financial Economics* 56, 3-28.

Chordia, Tarun, Richard Roll, and Avanidhar Subrahmanyam, 2002, Order imbalance, liquidity, and market returns, *Journal of Financial Economics* 65, 111-130.

Chordia, Tarun, and Avanidhar Subrahmanyam, 2004, Order imbalance and individual stock returns: Theory and evidence, *Journal of Financial Economics* 72, 485-518.

Coval, Joshua, and Erik Stafford, 2007, Asset fire sales (and purchases) in equity markets, *Journal of Financial Economics* 86, 479-512.

Daniel, Kent, and Tobias J. Moskowitz, 2016, Momentum crashes, *Journal of Financial Economics* 122, 221-247.

Dennis, Patrick J., and Strickland, Deon, 2002, Who Blinks in Volatile Markets, Individuals or Institutions?, *The Journal of Finance* 57, 1923-1949.

Fama, Eugene F., and French, Kenneth, R., 1992, The cross-section of expected stock returns, *The Journal of Finance* 47, 427-465.

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

Florackis, Chris, Andros Gregoriou, and Alexandros Kostakis, 2011, Trading frequency and asset pricing on the london stock exchange: Evidence from a new price impact ratio, *Journal of Banking & Finance* 35, 3335-3350.

Florackis, Chris, Alexandros Kontonikas, and Alexandros Kostakis, 2014, Stock market liquidity and macro-liquidity shocks: Evidence from the 2007–2009 financial crisis, *Journal of International Money and Finance* 44, 97-117.

Gârleanu, Nicolae, and Lasse Heje Pedersen, 2007, Liquidity and risk management, *American Economic Review* 97, 193-197.

Gârleanu, Nicolae, Lasse Heje Pedersen, and Allen M. Poteshman, 2009, Demand-based option pricing, *Review of Financial Studies* 22, 4259-4299.

Gorton, Gary, 2009, Information, liquidity, and the (ongoing) panic of 2007, *The American Economic Review* 99, 567-572.

Goyenko, Ruslan Y., Craig W. Holden, and Charles A. Trzcinka, 2009, Do liquidity measures measure liquidity?, *Journal of Financial Economics* 92, 153-181.

Gromb, Denis, and Dimitri Vayanos, 2002, Equilibrium and welfare in markets with financially constrained arbitrageurs, *Journal of Financial Economics* 66, 361-407.

Hameed, Allaudeen, Wenjin Kang, and S. Viswanathan, 2010, Stock market declines and liquidity, *The Journal of Finance* 65, 257-293.

Hasbrouck, Joel, 2009, Trading costs and returns for u.S. Equities: Estimating effective costs from daily data, *The Journal of Finance* 64, 1445-1477.

Hasbrouck, Joel, and Duane J. Seppi, 2001, Common factors in prices, order flows, and liquidity, *Journal of Financial Economics* 59, 383-411.

He, Zhiguo, and Arvind Krishnamurthy, 2013, Intermediary asset pricing, *American Economic Review* 103, 732-70.

Ho, Thomas S. Y., and Hans R. Stoll, 1983, The dynamics of dealer markets under competition, *The Journal of Finance* 38, 1053-1074.

Huberman, Gur, and Dominika Halka, 2001, Systematic liquidity, *Journal of Financial Research* 24, 161-178.

Jotikasthira Chotibhak, Lundblad Christian and Ramadorai Tarun, 2012, Asset fire sales and purchases and the international transmission of funding shocks, *The Journal of Finance* 67, 2015-2050.

Kyle, Albert S., 1985, Continuous auctions and insider trading, *Econometrica* 53, 1315-1335.

Lo, Andrew W., and Jiang Wang, 2006, Trading volume: Implications of an intertemporal capital asset pricing model, *The Journal of Finance* 61, 2805-2840.

Lo, AW, and J Wang, 2000, Trading volume: Definitions, data analysis, and implications of portfolio theory, *Review of Financial Studies* 13, 257-300.

Lou, Dong, 2012, A flow-based explanation for return predictability, *Review of Financial Studies* 25, 3457-3489.

Lou Xiaoxia, Sadka Ronnie, 2011, Liquidity level or liquidity risk? Evidence from the financial crisis, *Financial Analysts Journal* 67.

Ľuboš Pástor, and Robert F. Stambaugh, 2003, Liquidity risk and expected stock returns, *Journal of Political Economy* 111, 642-685.

Manconi, Alberto, Massimo Massa, and Ayako Yasuda, 2012, The role of institutional investors in propagating the crisis of 2007–2008, *Journal of Financial Economics* 104, 491-518.

Morris, Stephen, and Hyun Song Shin, 2004, Liquidity black holes, *Review of Finance* 8, 1-18.

Myron, S. Scholes, 2000, Crisis and risk management, *The American Economic Review* 90, 17-21.

Nagel, Stefan, 2012, Evaporating liquidity, *Review of Financial Studies* 25, 2005-2039.

R., Lynch Anthony W. and Mendenhall Richard, 1997, New evidence on stock price effects associated with changes in the s&p 500 index *Journal of Business* 70.

Shleifer, Andrei, 1986, Do demand curves for stocks slope down?, *The Journal of Finance* 41, 579-590.

Shleifer, Andrei, and Robert W. Vishny, 1997, The limits of arbitrage, *The Journal of Finance* 52, 35-55.

Spiegel, Matthew, and Avanidhar Subrahmanyam, 1995, On intraday risk premia, *The Journal of Finance* 50, 319-339.

Stoll, Hans R., 2000, Presidential address: Friction, *The Journal of Finance* 55, 1479-1514.

Vayanos, Dimitri, 2004, Flight to quality, flight to liquidity, and the pricing of risk, *Working Paper, NBER working paper series, London School of Economics*.

Vayanos, Dimitri, and Paul Woolley, 2013, An institutional theory of momentum and reversal, *Review of Financial Studies*.