Man vs. Machine: Quantitative and Discretionary Equity Management*

Simona Abis
Columbia University

July 31, 2017

Abstract

I use a machine learning technique to classify the universe of active US equity mutual funds as "quantitative", who mostly rely of computer-driven models and fixed rules, or "discretionary", who mostly rely on human judgement. I first document several new facts about quantitative investing. I then propose an equilibrium model in which quantitative funds have a greater information processing capacity but follow less flexible strategies than discretionary funds. The model predicts that quantitative funds specialize in stock picking, hold more stocks, display pro-cyclical performance, and that their trades are vulnerable to "overcrowding". In contrast, discretionary funds alternate between stock picking and market timing, display counter-cyclical performance and focus on stocks for which less overall information is available. My empirical evidence supports these predictions.

*Correspondence: Simona Abis, Columbia University, 3022 Broadway, Uris 420, New York, NY 10027. E-mail: sa3518@columbia.edu. I wish to thank Joël Peress for his support and invaluable advice. I am grateful for the very helpful comments from Lily Fang, Thierry Foucault, Sergei Glebkin, Denis Gromb, Marcin Kacperczyk, John Kuong, Hugues Langlois, Harry Mamaysky and Qi Zeng. I also thank Adrian Buss, Theodoros Evgeniou, Federico Gavazzoni, Naveen Ghondi, Massimo Massa, Ioanid Rosu, Lucie Tepla, Theo Vermaelen, Bart Yueshen, and Hong Zhang as well as seminar participants at INSEAD brownbags, the LabEx PhD consortium, the HEC PhD Workshop, and the INSEAD-Wharton Consortium and the AQR Institute Academic Symposium at LBS. I am responsible for any errors.

1 Introduction

Advances in computational technology and data analytics are revolutionizing the asset management industry (World Economic Forum 2015). As David Siegel, co-head of Two Sigma, a prominent hedge fund, puts it: "The challenge facing the investment world is that the human mind has not become any better than it was 100 years ago, and it's very hard for someone using traditional methods to juggle all the information of the global economy in their head (...). Eventually the time will come that no human investment manager will be able to beat the computer". Yet little is known about the prevalence of quantitative strategies in the mutual fund industry, their effect on fund performance, and their influence, if any, on asset prices. My paper fills this gap by carrying out a detailed analysis of quantitative investing in the US equity mutual fund universe.

The first step of my analysis is to develop a novel methodology to classify funds as quantitative or discretionary (i.e., non-quantitative). Here quantitative funds are those basing their investment processes primarily on quantitative signals generated by computer-driven models using fixed rules to analyze large datasets. Instead, discretionary funds rely mostly on decisions by asset managers who use information and their own judgment. I collect mutual fund prospectuses for 2,607 funds from 1999 to 2015 from the Securities and Exchange Commission (SEC). I manually categorize using objective criteria a subsample of 200 funds, which then serves as a training sample for a machine learning algorithm. That procedure results in a dataset of 599 quantitative and 1,851 discretionary funds.

The classification then allows me to report a number of stylized facts about quantitative funds. I find that they have quintupled in number and more than quadrupled in size over 1999–2015, growing at double the rate as discretionary funds. Currently their total assets under management (AuM) amount to $412 bn—or roughly 14% of the AuM of US equity mutual funds. Since 2011, in particular, capital has been flowing from discretionary to quantitative funds. On average, quantitative funds are younger (13.0 vs. 14.5 years) and smaller ($552 mm vs. $1.2 bn); they charge 10% lower expense ratios and 9% lower management fees, exhibit 10% higher turnover, own stocks that load more on value and momentum factors, and hold less cash than discretionary funds. They also experience outflows during recessions that are significantly stronger than those experienced by discretionary funds.[1]

I focus next on understanding how quantitative funds differ from their discretionary peers in terms of investment style and performance, and on how their growing importance might affect stock markets. One might expect quantitative funds to hold portfolios that are less susceptible to human managers' limited information processing capacity. On the other hand, these funds might be less flexible in adapting their strategies to changing market conditions, and their trades might be more correlated (Khandani and Lo 2007).

To guide my analysis of investment style and performance, I extend the model by Kacperczyk, Van

[1] All differences are significant at the 1% level.

Nieuwerburgh, and Veldkamp (2016; hereafter KVV). Theirs is a static general equilibrium model with multiple assets subject to a common aggregate shock and individual idiosyncratic shocks. These assets are traded by two categories of agents: skilled investors, and unskilled investors. Skilled investors have limited learning capacity, which they allocate to learning about idiosyncratic shocks (stock picking) or aggregate shocks (market timing). KVV show that skilled investors optimally focus their learning on the aggregate shock in recessions, modeled as periods of greater risk aversion and higher volatility of the aggregate shock[2], and on idiosyncratic shocks in expansions. The intuition is that the marginal benefit of learning about a shock increases with its volatility. So in recessions, when the aggregate shock volatility rises, the incentive to learn about it increases. This effect is magnified by the much larger supply of the aggregate shock (i.e., since it affects all stocks).

I interpret KVV's skilled investors as being discretionary investors. I then add to their model a second group of skilled investors: quantitative investors, whom I assume to have unlimited learning capacity but to only be able to learn about idiosyncratic shocks. This key assumption is motivated by survey evidence (Fabozzi, Focardi, and Jonas 2008) and by my textual analysis of prospectuses. Both pieces of evidence indicate that quantitative equity funds tend to specialize in stock-specific information and to incorporate marketwide dynamics in their models via price-based signals such as momentum.[3] Given the importance of this assumption for my analysis, I will further test its plausibility in the data.[4] The model yields seven testable predictions.

1. The presence of quantitative investors mitigates the pattern uncovered by KVV, whereby discretionary investors switch the focus of their learning across the business cycle. If quantitative investors represent a large enough share of the market, then discretionary investors instead choose to specialize in learning about the aggregate shock. This pattern reflects a substitution effect: the more investors learn about a given shock, the lower the value to others of learning about it.

2. Quantitative investors hold a larger number of stocks than discretionary investors. The reason is that investors optimally hold those assets about which their signals are more precise. Because of their unlimited capacity to learn about idiosyncratic shocks, quantitative investors have greater signal precision regarding the payoffs of more assets; hence it is optimal to hold more of them.

3. When learning about idiosyncratic shocks, discretionary investors focus on stocks for which relatively less information can be learned. That approach allows them to reduce their information disadvantage, or *information gap*, relative to quantitative investors.

[2]Various studies have shown that: aggregate stock market volatility is higher in recessions (Hamilton and Lin 1996; Campbell et al. 2001; Engle and Rangel 2008); risk premia and Sharpe ratios are counter-cyclical (Fama and French 1989; Cochrane 2006; Ludvigson and Ng 2009; Lettau and Ludvigson 2010); and aggregate risk aversion rises in recessions (Dumas 1989; Chan and Kogan 2002; Garlenau and Panageas 2012).

[3]Although this assumption is plausible for modeling equity mutual funds, it may be questionable in the context of some hedge funds (e.g. quantitative macro funds specialize in timing the market using trend-following signals and signals derived from the analysis of macroeconomic information).

[4]I report below two findings supporting this hypothesis. First, I find that quantitative funds that display consistently high stock-picking ability in expansions do not switch to displaying high market-timing ability in recessions, indicating that these funds indeed are not flexible in changing the focus of their strategies. Second, I find that that no funds consistently display high market-timing ability throughout the business cycle, indicating in turn that no fund is endowed with the ability to learn only about aggregate shocks.

4. Dispersion of opinion, and hence of holdings, is greater among discretionary investors than among quantitative investors. This follows from discretionary investors' limited learning capacity which they need to allocate optimally across shocks. They might choose to learn about different shocks, as the marginal benefit of learning about shocks is decreasing in the average precision of private signals about those shocks in the market – substitution effect. This leads to more disperse portfolio allocations.

5. Quantitative investors display pro-cyclical performance. In expansions, though, two counteracting effects are at play: higher signal precision about idiosyncratic shocks (which, in expansions, matter more than the aggregate shock) helps quantitative investors outperform the market; but if the fraction and signal precision of quantitative investors are too high, their advantage is eroded and they tend to underperform.

6. Performance is counter-cyclical for discretionary investors (as in KVV) thanks to their superior information about the aggregate shock, which is more relevant in recessions. As the share of quantitative investors rises, their performance is worsened in expansions and improved in recessions, due to a substitution effect.

7. As the share of quantitative investors increases, so does the price informativeness of idiosyncratic shocks. However, the effect on the aggregate shock's price informativeness varies over the business cycle. In expansions, aggregate price informativeness is enhanced but only when the share of quantitative investors is high enough that discretionary investors are induced to specialize in learning about the aggregate shock (Prediction 2). In recessions, aggregate price informativeness always decreases with the share of quantitative investors; this is because an increase in the share of quantitative investors reduces that of discretionary investors, who are the only investors able to learn about the aggregate shock.

To test these predictions, I merge my sample with the CRSP mutual fund and stock databases, the Thompson Financial Spectrum dataset, the CRSP Mutual Fund Holdings dataset, IBES analysts forecasts, and Dow Jones news.

First, I study the extent to which funds engage in stock picking and market timing measured, as in KVV, by the covariance of fund holdings with future earnings surprises and with future growth in industrial production respectively.

To begin with, I find that KVV's prediction applies only to funds I classify as discretionary: such funds that display consistently high stock-picking ability in expansions also show high market-timing ability in recessions (68% higher than the average fund).

Next I test the plausibility of my new key assumption—that is, quantitative funds only learn about idiosyncratic shocks. Quantitative funds that display consistently high stock-picking ability in expansions do not display high market timing ability in recessions, and no funds consistently display high market-timing ability throughout the business cycle. These findings lend support to the model's assumption about funds' learning technologies by suggesting, respectively, that quantitative funds do not flexibly adapt the focus of their learning throughout the business cycle, and that no fund is endowed with the ability to learn only about aggregate shocks.

Finally, in accordance with Prediction 1, I find that, with increases in the share of quantitative funds (measured as the relative TNA that quantitative funds manage or as the fraction of the US stock market capitalization held by quantitative funds in my sample), the stock-picking ability of discretionary funds in expansions decreases significantly (by one standard deviation for every 17% increase in the share of quantitative funds) and their market-timing ability in recessions increases significantly (by one standard deviation for every 6.8% increase in the share of quantitative funds).

Second, I find that quantitative funds on average hold considerably more stocks than do discretionary funds (225 vs. 117). Additionally, the distribution of quantitative funds' number of holdings is more skewed, with a significant portion of funds holding up to 1000 stocks. This translates into greater portfolio diversification, and hence lower idiosyncratic volatility (up to 15% less). This result is in line with Prediction 2.

Third, I proxy for the average information gap in the stocks held by quantitative and discretionary funds with their average size, age, media mentions and analysts coverage. The idea is that more information is available about stocks for which there exists a longer history and that enjoy greater media and analyst coverage, making such stocks more easily processable with quantitative approaches; hence the information gap should be greater. I find that the information gap is smaller for the stocks held by discretionary funds; they hold stocks that are younger (44 months younger than the 351 months average) and have fewer monthly mentions in the media (31 fewer than the 275 average).[5] Although there is no significant difference in analyst coverage, I find that all funds holding stocks less mentioned in the news or less covered by analysts perform marginally better—and that the effect is significantly greater for discretionary funds (about 36% higher depending on the measure). This result is consistent with Prediction 3.

Fourth I find that discretionary investors have significantly more disperse holdings than quantitative funds do. This result holds for different measures of dispersion of holdings: the cumulative squared difference in the weight allocated by each fund to stocks relative to the weight allocated by the average fund of the same type (dispersion), and the average percentage of funds of the same type that hold the same stocks (commonality). This result is in line with Prediction 4.

Fifth, I find that the performance of quantitative funds in expansions has been declining over time due to "overcrowding", caused by the increase in the share of quantitative funds in the market and their tendency to hold overlapping portfolios. In the post-subprime expansion (Jul 2009–Dec 2015), quantitative funds earn 79.8 bp less than before the 2001 recession (Dec 1999–Feb 2001) and 12 bp less than in the expansion that preceded the sub-prime crisis (Jan 2002–Nov 2007). As a consequence, in the latest expansion they do not significantly outperform the market, despite displaying better Sharpe ratios. Finally, funds with a lower average commonality in holdings display better performance. This outcome is in line with Prediction 5.

[5] Differences are statistically significant.

Sixth, I find that discretionary funds' performance is counter-cyclical: they perform significantly better in recessions (by 31.0 bp for CAPM alphas, 8.5 bp for three-factor alphas and 2.9 bp for 4-factor alphas); while displaying a small negative alpha in expansions (insignificant for CAPM alphas, -7 bp for 3-factor alphas and -5 bp for 4-factor alphas). Such behavior is consistent with Prediction 6.

The paper proceeds as follows. Section 2 reviews the related literature. Section 3 describes the classification method and new facts about quantitative and discretionary funds. Section 4 presents the model and its implications. Section 5 reports empirical tests and robustness checks. Section 6 concludes. The Appendix features details of the classification method and the proofs of the model.

2 Related Literature

My paper relates to three streams of research. First and foremost, it contributes to the literature on mutual funds. Research on quantitative investing therein is scarce and relies on small samples because of the difficulty in identifying quantitative traders. Casey, Quirk & Associates (2005) report that quantitative managers outperform discretionary funds (32 funds over 2000–2004), and Ahmed and Nanda (2005) that they outperform their peers in small-cap stocks and charge lower management fees (22 funds over 1980–2000). Khandani and Lo (2007) attribute the August 2007 "Quant Meltdown"–sudden losses suffered by many long-short quantitative hedge funds–to the rapid unwinding of "overcrowded" strategies. Finally Harvey et al. (2016) analyze differences in performance between quantitative and discretionary hedge funds. They find that discretionary and quantitative equity hedge funds perform similarly while quantitative macro hedge funds outperform their discretionary counterparts. I find, in my comprehensive sample, somewhat consistent results (e.g., on fees, performance and portfolio overlap), but show that many need to be qualified, for instance to account for the rise in quantitative investing and for business cycles.

Additional evidence comes from the survey by Fabozzi, Focardi, and Jonas (2008). The survey underscores the diminished relative performance of quantitative funds during the subprime crisis, for which some of the main causes were identified as rising correlations and fundamental market shifts. Respondents indicated that adding macroeconomic concepts directly into models was one of the least useful potential improvements–an expression of how the successful forecasting of financial markets relies on model adaptability (Fabozzi and Focardi 2009). Yet most contemporary adaptive models (e.g., regime-switching models) require very long time series for accurate estimation, which are generally not available for financial data. I find confirmation of this preliminary evidence in the analysis of the prospectuses of the funds I classify. This motivated my choice of assumptions in modeling quantitative funds' learning technology.

An important branch of the mutual fund literature focuses on performance. It offers little evidence that mutual funds outperform the market unconditionally, be it by picking stocks or timing the market. Conducting a finer analysis, recent studies demonstrate that their performance varies over the business cycle (Glode 2011; Kosowski 2011; de Souza and Lynch 2012). Kacperczyk, Van Nieuwerburgh, and Veldkamp (2014), in particular, show empirically that managers who are good stock-pickers in expansions are also good market-timers in recessions. My empirical results underscore too the importance of conditioning on the state of the business cycle, and of distinguishing market timing from stock picking.

From a theoretical perspective, the shifting skills and performance of funds across the business cycle has been explained through rational inattention (Sims 2003; Mackowiak and Wiederholt 2009 and 2015 ; Van Nieuwerburgh and Veldkamp 2010; Kacperczyk, Van Nieuwerburgh, and Veldkamp 2014; KVV). This literature argues that investors with limited information processing capacity make rational choices about what information to focus on. KVV show both theoretically and empirically that these investors optimally shift their attention from idiosyncratic shocks in expansions to aggregate shocks in recessions, leading to cyclicality in performance. My model extends theirs by allowing skilled investors with different learning capabilities and flexibilities to interact in equilibrium. Thus, I explain how quantitative and discretionary investors compete in equilibrium and how the rise in quantitative investment affects the choices and profitability of discretionary funds. Glasserman and Mamaysky (2016) also contribute to this debate by developing a model in which some investors are informed about aggregate shocks and others about idiosyncratic shocks, but only the latter trade against idiosyncratic noise traders. They show that an endogenous bias exists towards idiosyncratic informativeness. This is consistent with my model when the share of quantitative investors is large.

The paper's second main contribution is to the growing literature that studies the impact of technological changes on financial markets. The themes therein include payments systems (e.g., bitcoin), disintermediation (e.g., peer-to-peer lending), and disruptive innovations (e.g., fintech). Most work on quantitative investing focuses on specific investors, high-frequency traders ("HFT"), who exploit progress in communications technology and machine-based analytics to execute trades at very high speed. Here the unit of analysis is a trade, not a trader, thus circumventing the aforementioned identification challenge. Though worries about adverse selection remain important, most of the evidence suggests that HFT improve market quality through lower bid-ask spreads (Menkveld 2015, Bershova and Rakhlin 2013, Malinova, Park, and Riordan 2013, Stoll 2014, Riordan and Storkenmaier 2012, Boehmer, Fong, and Wu 2015) and more efficient pricing (Carrion 2013, Riordan 2014a, Conrad, Wahal, and Xiang 2015, Chaboud et al. 2014, Hasbrouck and Saar 2013). An ongoing criticism is that HFT exacerbate market movement and magnify volatility because their trades tend to be highly correlated (Shkilko and Sokolov 2016). Though my paper does not speak directly to the literature on HFT (in my model, all trading takes place at the same speed), I uncover in the data a trade-off which has a similar flavor. Specifically, I find, on one hand, that quantitative investing contributes to price efficiency, but on the other, that it has the potential to magnify market instability because machine-based trades tend to overlap significantly.

The third stream of literature related to my paper analyses stocks connectivity and fragility. Recent research shows that stocks with greater commonality in their ownership are more exposed to non-fundamental risk. Greenwood and Thesmar (2011) define stocks as fragile if they are exposed to more non-fundamental risk either because of concentrated ownership or because owners face correlated liquidity shocks; finding that fragility predicts stocks volatility. Anton and Polk (2014) show that the degree of shared ownership forecasts cross-sectional variation in return correlation. I find evidence that quantitative funds have greater commonality in their portfolio holdings. The rise in these types of funds might in turn impact stock return volatility and correlation.

3 Quantitative versus Discretionary Funds

I develop a novel methodology, based on machine-learning analysis of fund prospectuses, to classify funds as quantitative or discretionary. In this section I describe the data employed and also the classification methodology. I then report a series of new facts about quantitative and discretionary funds based on this classification.

3.1 Data

Following the path of most data availability, I focus on diversified active US equity mutual funds featured in the CRSP Mutual Fund dataset. Thus I exclude international funds, sector funds, unbalanced funds, index funds, and underlying variable annuities. Incubation bias is eliminated by focusing only on those observations that are dated *after* the fund inception date (Elton, Gruber, and Blake 2001). To focus on funds for which holdings data are most complete, I also exclude funds with less than $5 million AuM or holding fewer than 10 stocks or devoting less than 80% of their portfolios to standard equities (Kacperczyk, Sialm, and Zheng 2008).

Funds often market different share classes of the same portfolio differing only in target clients, fees structure or fund withdrawal restrictions. I follow the mutual fund literature and group the observations of all share classes of a given fund for every point in time into a single observation. Toward that end I keep the qualitative variables (e.g., fund name, start date) of the oldest fund, add the TNA of all share classes, and weight all other variables (e.g., returns) by lagged total net assets.

Stock holdings data are obtained from merging the Thompson CDA/Spectrum Mutual Fund Holdings dataset for 1999–2003 and the CRSP Mutual Fund Holdings dataset for 2004–2015 to the CRSP Survivorship-Bias-Free mutual fund dataset. This choice is dictated by the fact that the CRSP dataset also reports short positions held by mutual funds, while Thompson doesn't. Different short-selling practices are one of the characteristics along which quantitative and discretionary funds differ. CRSP holdings are available starting from July 2001 but the holding's effective dates is only reported starting from December 2003, which is when I start using it. The sample period starts in December 1999 and ends in December 2015; that is the period for which fund prospectuses contain the information needed for my classification (see Section 3.2). This final condition restricts the sample to 2,607 unique funds and 250,427 fund-month observations. The number of funds each month ranges from 516 funds in (December 2000) to 2,065 funds (March 2011); see Figure 1.

To construct different variables, I also use the CRSP stock-level database, seasonally adjusted industrial production from Federal Reserve economic data, Compustat earnings, IBES forecasts, Fama–French factors, and Dow Jones news mentions.

3.2 Classification Methodology

I collected prospectuses for the 2,607 funds from December 1999 to December 2015 from the SEC online database. One section of the fund prospectus, usually entitled "Principal Investment Strategies", corresponds to item 9 of the N-1A mandatory disclosure form, where funds must disclose "The Fund's principal investment strategies, including the particular type or types of securities in which the Fund principally invests or will invest" and "Explain in general terms how the Fund's adviser decides which securities to buy and sell." This requirement was added after the 1998 amendment of mandatory disclosures; all funds were required to comply beginning in December 1999, the starting date of my analysis.

Quantitative funds are defined as those with an investment process based primarily on quantitative signals generated by computer-driven models using fixed rules to analyze large datasets. *Discretionary* funds are defined as those with an investment process based mainly on decisions made by asset managers, who use information and their own judgment. Although discretionary funds may well employ quantitative measures, their main stock selection criterion remains human judgment. Conversely, quantitative funds might incorporate some level of human overlay, for instance, double-checking the soundness of automatically generated trading signals. Hybrid cases, in which extensive quantitative stock screening is used before discretionary analysis determines the final trading decision, are grouped with quantitative funds.

Most funds indicate, in their "Principal Investment Strategies", whether they employ quantitative methods. For quantitative funds, a typical statement is as follows:

> The Portfolio's subadviser utilizes a quantitative investment process. A quantitative investment process is a systematic method of evaluating securities and other assets by analyzing a variety of data through the use of models—or systematic processes—to generate an investment opinion. The models consider a wide range of indicators—including traditional valuation measures and momentum indicators. These diverse sets of inputs, combined with a proprietary signal construction methodology, optimization process, and trading technology, are important elements in the investment process. Signals are motivated by fundamental economic insights and systematic implementation of those ideas leads to a better long-term investment process. (Advanced Series Trust AST AQR Large-Cap Portfolio)

In contrast, a discretionary fund's prospectus might state:

> The fund's portfolio construction combines a fundamental, bottom-up research process with macro insights and risk management. The fund's portfolio managers, supported by a team of research analysts, use a disciplined opportunistic investment approach to identify stocks of companies that the portfolio managers believe are trading materially below their intrinsic market value, have strong or improving fundamentals and have a revaluation catalyst. The fund seeks exposure to stocks and sectors that the fund's portfolio managers perceive to be attractive from a valuation and fundamental standpoint. Portfolio position sizes and sector weightings reflect the collaborative investment process among the fund's portfolio managers and research analysts. The portfolio managers also assess and manage the overall risk profile of the fund's portfolio. (Advantage Funds Inc.: Dreyfus Opportunistic US Stock Fund)

I classified manually a subsample of 200 prospectuses from different funds while using various objective criteria: the word "quantitative" or "systematic" appearing in the fund name, the fund being identified in the media as a quantitative or discretionary mutual fund. Next, I pre-processed the prospectus extracts describing the firm's principal investment strategies and employed the "bag of words" approach to transform that text into a matrix of *features* (i.e. stemmed words and two-words combinations) suitable for automatic processing.[6]

I then used nested cross-validation to train different machine learning classification algorithms: bernoulli and multinomial naive Bayes, logistic regression, linear and nonlinear support vector machines, k-nearest neighbors, and random forest. The goodness of fit is measured as the accuracy in classifying a validation sample *divided by* the standard deviation in validation accuracy. I chose the most transparent and accurate algorithm to classify the rest of the population.

The chosen algorithm is the random forest with an ensemble of 1,000 trees and an entropy-based impurity measure. I finally ran the estimated model on a sample that had been classified but not yet used, obtaining an out-of-sample accuracy of 93.4%. The random forest is a relatively transparent algorithm in that it allows one to visualize the features being utilized along with their relative informational content, measured as the reduction in entropy obtained with each specific feature. This algorithm has the additional benefit of yielding not only a binary classification but also a probability that a given fund belongs to a category. [7]

[6]In pre-processing, stop-words (is, the, and, etc.) and financial stopwords (S&P, Russel, ADR, etc.) were eliminated and words were stemmed to their root using the Porter stemmer algorithm (thus, e.g.: quantitative, quantitatively, ... = quantit). I then compiled a comprehensive list of all features mentioned in the sample. All features that appeared in more than 60% (or in less than 5%) of the files were eliminated because unlikely to be informative in distinguishing the two fund categories. Next, the stemmed features were converted into feature vectors, which contain a 1 (resp., 0) when the focal text file does (resp., does not) mention the feature. Finally, the features weights were adjusted by within-text term frequency versus the frequency in all samples (tf-idf).

[7]In the binary classification a fund is considered to be quantitative (resp. discretionary) when the probability of being quantitative (resp. discretionary) is greater than 50%.

3.3 New Basic Facts about Quantitative and Discretionary Funds

I use my classification to document some new facts about quantitative and discretionary funds.

A first series of facts relates to the prospectuses themselves. The algorithm uses 828 features, but the first 10 account for 21% of the informativeness with respect to reducing classification impurity (as evaluated using an entropy measure; see Figure 2). Most features are either intuitive or easily interpretable when bearing in mind the algorithm's conditional nature.[8] The most discriminating features are the words "quantitative", "proprietary" and "model". It is interesting that one of the top 10 features distinguishing quantitative from discretionary funds is the word "momentum". This finding is in line with survey evidence that quantitative funds mostly use momentum, sentiment, and other trend-based measures to incorporate market dynamics into their models (Fabozzi, Focardi, and Jonas 2007).

A word search of the classified documents (see Appendix A) shows that, relative to discretionary funds, quantitative fund prospectuses refer more often to active trading (8.5% vs. 3.8%), frequent trading (6.8% vs. 3.8%), and short selling (9.0% vs. 3.8%). This last accords with one of the reasons commonly associated with the rise of quantitative mutual funds: their better ability to incorporate strategies that involve long–short trades, such as the "130–30" strategy (Fabozzi, Focardi, and Jonas 2007).[9] It is noteworthy that a much higher proportion of quantitative funds report following such strategies as momentum, sentiment, and technical analysis (30.3% vs. 5.0% for discretionary funds), a result that is also in line with the survey evidence already cited. Overall, about 25% of all funds tracked are quantitative.

I uncover another set of facts by merging the sample with the CRSP Survivorship-Bias-Free dataset. Thus I am able to quantify the rise of quantitative funds. Over 1999–2015, quantitative funds quintupled in number (from 94 to 465 out of a 2015 industry total of 1,956) and quadrupled in AuM (from $99 bn to $412 bn out of a $3 tn total). During this period, quantitative funds grew nearly twice as fast as discretionary funds: from 8.8% to 14% of the market's total net assets (TNA); see Figure 3.

The differences in quantitative and discretionary funds were explored along six dimensions: age, size, holdings, shortselling, fund flows, and fee structure.

Quantitative funds are, on average, both younger (13.0 vs. 14.5 years) and smaller ($552 million vs $1.2 billion TNA) than discretionary funds, the age difference is less pronounced in recessions. The

[8] For instance, one of the most relevant features is "based". Although "based" is not in itself that informative, conditional on the presence of words such as "quantitative" or "quantitative model" it increases the likelihood of a fund being quantitative. Many quantitative funds describe their strategy in their prospectuses with sentences of the following type: "The fund uses a quantitative model based on…". A few features remain difficult to interpret. A possible reason is that, when features are highly correlated, only one of them is given importance by the algorithm—using criteria that do not include interpretability.

[9] A 130–30 strategy consists of 130% long exposure, together with 30% short exposure, so that *net* exposure to the market is 100% (i.e., long minus short) while *gross* exposure is 160% (long plus short). Commonly used analogues include the 110–10, 120–20, and 150–50 strategies.

effects are significant at the 1% level even after controlling for various fund-level characteristics (see Table 1).

The age difference has two sources. First, quantitative funds began to appear later than did discretionary funds. As shown in Figure 4, the relative age differential has therefore been (mechanically) declining simply because of time passing. Second, the creation and failure rates of discretionary and quantitative funds differ significantly. Figure 5 shows that, until 2009, quantitative funds enjoyed a significantly higher creation rate and a lower failure rate than did discretionary funds. This fact explains the steep rise in the relative number of quantitative funds displayed in Figure 3 for that period. In fact none of the quantitative funds in my sample failed during 2003–2007 and few did even in 2008 and 2009. This may explain the slightly smaller age gap in recessions, as reported in Table 1. Since 2009, however, the creation rate of both discretionary and quantitative funds has declined significantly while their failure rates have risen, thereby stabilizing the total number of funds.

Quantitative funds are smaller (in terms of TNA) than discretionary funds (Figure 6). Whereas the size distribution of discretionary funds is approximately log-normal and extremely right-skewed, quantitative funds exhibit a higher concentration of very small funds and a distribution that is less skewed to the right. To study formally the size distributions, I ran a quantile regression of fund size on quantitative and recession indicator variables and on various other fund-level characteristics:

$$TNA_{jt} = \alpha^p + \beta_1^p NBER_t + \beta_2^p Quant_j + \beta_3^p Quant_j \times NBER_t + \gamma^p X_{jt} + \varepsilon_t.$$

Here $NBER_t$ is a dummy set to 1 only for recessions and $Quant_t$ is a dummy set equal to 1 only for quantitative funds; TNA_{jt} are the total net assets held by fund j at time t; and the X_{jt} are demeaned, fund-specific control variables. Finally the superscript "p" indicates the different percentiles of the size distribution for which the regression is run. Detrending allows interpreting α^p as the size of discretionary funds and β_2^p as the difference in size between quantitative and discretionary funds (β_2^p/α^p represent percentage size difference).

I find that the relative negative effect of the quant dummy on fund size ($\beta_2^p/\alpha^p < 0$) decreases for higher size percentiles. The ratio is significantly different for most percentile combinations at the 1% level, with the exception of a few percentiles in the mid-section of the size distribution. Figure 8 illustrates the quantile regression results graphically. At the 1^{st} percentile of the size distribution, quantitative funds are 38% smaller than discretionary funds; at the 99th percentile, they are only 16% smaller. Hence quantitative funds—despite generally being smaller than discretionary funds—display a stronger tendency toward polarization between extremely small (boutique) funds and larger ones, where the latter are closer to the size of their discretionary counterparts. These plotted values explain why quantitative funds represented 24.4% of the total number of funds in the market yet managed only 14% of the TNAs.

I uncover the third main set of differences by analyzing fund holdings. Portfolio holdings turnover is about 10% faster for quantitative than for discretionary funds. Measuring fund holdings' exposure to the momentum, "small [market cap.] minus big" (SMB), and "high [book/market] minus low" (HML) factors as the TNA-weighted average of the stock-level load on those factors, I find that quantitative funds hold stocks that load more on the momentum factor during expansions but holding stocks with more *negative* momentum in recessions (contrarian). These funds are not exposed to the SMB factor. In contrast, discretionary funds load positively on the size factor (i.e., small stocks). In comparison with discretionary funds, quantitative funds load more on the HML factor (i.e., value stocks). Quantitative funds hold 23% less cash and do not increase their cash holdings in recessions, whereas discretionary funds increase their cash holdings by 16% during recessions (Table 2).

Fourth there is some evidence that quantitative funds shortsell more. The CRSP holdings dataset, which starts in December 2003, also reports short positions. In my sample short positions are available for 54 funds of which 30 discretionary and 24 quantitative. Quantitative funds short-sell on average 16% of gross TNA (TNA of short + long positions), as opposed to 5.4% for discretionary fund. Figure 7 shows the average percentage of gross TNA sold short by quantitative and discretionary funds over time. We observe that quantitative funds short-sell a much larger percentage of their portfolios and this amount has been increasing since 2011.

Fifth, quantitative funds experience significant outflows in recessions (measured as the monthly percentage change in TNA not due to returns), while fund flow volatility is not significantly different between quantitative and discretionary funds (Table 3). Figure 9 depicts the evolution of the mean and median fund flows from 1999 to 2015. We can see a significant difference in the behavior of fund flows before and after the subprime crisis; more specifically, the median fund has been experiencing significant outflows since 2007. We also observe that, in line with Table 3, average outflows during both the new millennium's recession and the subprime crisis that followed are greater from quantitative than from discretionary funds. For the last five years of the study, though, we observe a progressive shift of capital from discretionary to quantitative funds, with the average discretionary fund experiencing outflows and the average quantitative fund enjoying inflows. This is both evident graphically and statistically significant, as shown in Model 3 of Table 3. This is the period of steepest relative growth of quantitative funds, as shown in Figure 3.

Finally, quantitative funds charge an 11% lower expense ratio. This ratio is the total fee charged annually for fund expenses as a percentage of assets. Examining some of the expense ratio's components reveals that management fees are significantly (9%) lower for quantitative funds while the "12b1" fee, which includes marketing and distribution fees, exhibits no statistically significant difference by fund type. Quantitative and discretionary funds do not charge different fund loads (Table 4).

4 A Model of Financial Markets with Quantitative and Discretionary Funds

In this section I describe a theory of how quantitative and discretionary funds differ in their investment policy and performance—and of how these differences vary over the business cycle. The model builds on Kacperczyk, Van Nieuwerburgh, and Veldkamp (2016). Theirs is a static general equilibrium model with multiple assets subject to a common aggregate shock and to idiosyncratic shocks. The assets are traded by skilled investors, unskilled investors, and noise traders. Skilled investors can learn about assets' payoffs, but their learning capacity is limited. I augment that model by adding a second group of skilled investors, quantitative investors, endowed with an unlimited learning capacity but able to learn only about idiosyncratic shocks. Quantitative investors aside, the only difference between this model and KVV's is my assumption that private signals contain an unlearnable component (i.e., residual noise) that is heterogeneous across assets.

4.1 Model

The model has three dates, $t = 1, 2, 3$. At $t = 1$, investors allocate their learning capacity. At $t = 2$, investors choose their portfolio allocations. At $t = 3$, prices and returns are realized.

4.1.1 *Assets and Risk Factors*

There are n risky assets and one riskless asset, with price of 1 and payoff r. Of the risky assets, $n-1$ are exposed to both idiosyncratic and aggregate risks, while n is a composite asset subject to the aggregate risk only. The risky assets' payoffs at $t = 3$ are normally distributed and given by:

$$f_i = \mu_i + b_i z_n + z_i \text{ and } f_n = \mu_n + z_n. \tag{1}$$

Here f_i is the payoff from asset i, $i = 1, \ldots, n$; the risk factors are given by $z = [z_1, \ldots, z_n]' \sim N(0, \Sigma)$; where z_n represents the aggregate shock and z_i for $i \neq n$ represents idiosyncratic shocks. Σ is a diagonal matrix such that $\Sigma_{ii} = \sigma_i \in \mathbb{R}_+$. The term b_i is asset i's exposure to the aggregate risk, and $\mu_i \in \mathbb{R}$ is its expected payoff. Rewriting the system (4.1) in matrix form yields $f = \mu + \Gamma z$. The model is solved in terms of "synthetic" payoffs, which are affected by only one risk factor each: $\tilde{f} = \Gamma^{-1} f = \Gamma^{-1}\mu + z$.

Since the supply of risk factors is stochastic, prices are not fully revealing. For $i = 1, \ldots, n$, the supply of the ith risk factor is $\bar{x}_i + x_i$, where $x_i \sim N(0, \sigma_x)$ and the overbar signifies the expected component of supply. One important assumption is that the aggregate risk factor is in much greater supply than any other risk factor because it affects all n risky assets, so $\bar{x}_n \gg \bar{x}_i$ for $i \neq n$.

4.1.2 *Investors and Learning*

There is a unit mass of mean-variance investors with risk aversion ρ, indexed by $j \in [0, 1]$, of whom a fraction $\chi \in [0, 1]$ are skilled, the rest being unskilled. Among skilled investors, a fraction $\theta \in [0, 1]$ are quantitative and $(1-\theta)$ are discretionary investors. Thus the measures of quantitative and discretionary investors are, respectively, $\theta\chi$ and $(1-\theta)\chi$.

Investors receive private signals about risk factors. The greater their capacity for learning, the greater the signals' precision. Investor j's private signals vector is $s_j = [s_{1j}, \dots, s_{nj}]'$ such that:

$$s_{ij} = z_i + \epsilon_{ij},\ where:\ \epsilon_{ij} \sim N(0, \sigma_{ij})\ for\ \sigma_{ij} \in [\underline{\sigma}_{ij}, \infty] \tag{2}$$

The capacity for learning determines by how much investor j is able to reduce the volatility of the private signal noise (σ_{ij}). An infinite learning capacity allows to reduce the error term to the lower-bound $\underline{\sigma}_{ij}$; whereas zero learning capacity translates into an infinite variance of the private signal noise.

Discretionary investors ($j = d \in [0, \chi(1-\theta)]$) have a total learning capacity K, which they can allocate freely across all risk factors; so that the sum of their signals' precisions is bounded:

$$\sum_{i=1}^{n} \sigma_{id}^{-1} = K\ where\ \sigma_{id}^{-1} \geq 0\ \forall i = 1, \dots n \tag{3}$$

Q*uantitative* investors ($j = q \in [0, \chi\theta]$) have unlimited capacity for learning about idiosyncratic risk factors for all $n = 1, \dots, n-1$. But they do not receive a private signal about the aggregate shock $\sigma_{nq} = \infty$.

Unskilled investors ($j = u \in [0, 1-\chi]$) have zero learning capacity hence $\sigma_{ij} = \infty\ \forall i = 1, \dots n$; which is equivalent to saying that they do not receive any private signal.

All investors, skilled or unskilled, learn from prices through the signals vector $s_p = [s_{1p}, \dots, s_{np}]'$, where:

$$s_i^p = z_i + \epsilon_{ip},\ where\ \epsilon_{ip} \sim N(0, \sigma_p). \tag{4}$$

4.1.3 *Recessions*

I derive differential predictions for expansions and recessions. Recessions are characterized empirically as periods in which both the volatility of the aggregate shock and the price of risk rise significantly (see the works cited previously in footnote 2). Following KVV, I model recessions as periods of higher aggregate shock volatility (σ_n) and higher risk aversion than expansions.

4.2 Analysis

4.2.1 *Optimal Portfolio Choice*

In period $t = 2$, each investor j ($j = u, d, q$), given initial wealth W_0 and having risk aversion ρ, chooses the optimal portfolio allocation $\tilde{q}_j^*$ to maximize a mean-variance utility function:

$$\max_{\tilde{q_j}} U_{2j} = \left\{ \rho E_j[W_j] - \frac{\rho}{2} V_j[W_j] \right\} \text{ s.t.} W_j = rW_0 + \tilde{q_j}'(\tilde{f} - \tilde{p}r). \tag{5}$$

It follows that

$$\tilde{q_j}^* = \frac{1}{\rho} \hat{\Sigma_j}^{-1} \left(E_j[\tilde{f}] - \tilde{p}r \right). \tag{6}$$

The allocation to risky assets decreases with risk aversion ρ, but increases with the posterior private signal's precision $\hat{\Sigma_j}^{-1}$ and with the expected payoff $E_j[\tilde{f}]$; note that the last two measures are group dependent. Skilled investors, who have more precise signals, optimally allocate more of their capital to risky assets.

4.2.2 *Market Clearing*

Given the optimal portfolio choices of the different investors, the next step is to clear the asset market by equating the aggregate demand to supply such that:

$$\int \tilde{q}_j^* \, dj = \bar{x} + x. \tag{7}$$

The solution to the integral in equation (7) depends on the average learning capacity, across investor classes, toward the different risk factors:

$$\bar{K}_i = \int K_{ij} \, dj = \begin{cases} \chi\theta K_{iq} + \chi(1-\theta)K_{id}, & i = 1, ..., n-1; \\ \chi(1-\theta)K_{id}, & i = n; \end{cases} \tag{8}$$

where $K_{iq} = \int \sigma_{iq}^{-1} \partial q$ and $K_{id} = \int \sigma_{id}^{-1} \partial d$.

I solve for an equilibrium price of the form $pr = (A + B + Cx)$, where (A, B, C) depend on the parameters.

4.2.3 *Investors' Learning Choice*

At $t = 1$, all skilled investors choose their optimal learning capacity allocation.

Quantitative investors, being unlimited in their learning capacity, learn all available and machine processable information about idiosyncratic shocks such that $\sigma_{iq}^{-1} = \underline{\sigma}_{iq}^{-1}$ $\forall i = 1, ..., n-1$ and $K_{iq} = \underline{\sigma}_{iq}^{-1}$.

Discretionary investors, having a limited learning capacity, must optimize its allocation. This choice depends on investors' expectation, at $t = 1$, of the distribution of excess returns at $t = 2$. After some manipulation, their expected utility at $t = 1$ can be written as:

$$U_{1d} = \frac{1}{2} \sum_{i=1}^{n} (\sigma_{id}^{-1} \lambda_i) + \text{constant} \tag{9}$$

for λ_i the marginal benefit of learning about risk factor i:

$$\lambda_i \equiv \bar{\sigma}_i\{1 + [\rho^2(\sigma_x + \bar{x}_i{}^2) + \bar{K}_i]\bar{\sigma}_i\}, \tag{10}$$

$$\bar{\sigma}_i = \int \left(\hat{\Sigma}_j\right)_{ii} dj = \left(\sigma_i^{-1} + \bar{K}_i + \frac{\bar{K}_i{}^2}{\rho^2\sigma_x}\right)^{-1}. \tag{11}$$

According to equation (10), λ_i increases with expected supply $\bar{x}_i$ and prior volatility σ_i but decreases with average private information in the market, $\bar{K}_i$. The latter result is due to a substitution effect: when many investors learn about a given shock, the benefit that each derives from that knowledge is reduced. Equation (8) shows that $\bar{K}_i$ is a function of the share χ of skilled investors among all investors and of the fraction θ, among those skilled investors, that are quantitative.

The learning problem of discretionary skilled investors is given by:

$$\begin{gathered} \max_{K_{1d}\ldots K_{nd}} a\tfrac{1}{2}\textstyle\sum_{i=1}^{n}(\sigma_{id}^{-1}\lambda_i) + \text{constant} \\ \text{s.t.} \textstyle\sum_{i=1}^{n}\sigma_{id}^{-1} \leqslant K, \\ \underline{\sigma}_{id}^{-1} \geq \sigma_{id}^{-1} \geq 0 \quad \forall i = 1, ..., n. \end{gathered} \tag{12}$$

The solution to this problem consists in allocating all learning capacity to the risk factor i^* (when $i^* = \text{argmax}_i\, \lambda_i$) or to a basket of risk factors i^* such that $\lambda_{i^*} \in I_M/i^*$ and $\lambda_{i^*} = \text{argmax}_i\, \lambda_i$ (Van Nieuwerburgh and Veldkamp 2010). It is possible to have a basket of risky factors that yield the same marginal benefit of learning because λ_i^* is itself also a function of average learning; hence substitution effects result in the increase in average precision about a risk factor reducing its appeal.[10]

An important difference from the KVV approach is that here the average precision of the signal about each risk factor ($\bar{K}_i$) also depends on the learning advantage (or disadvantage) of the quantitative investors—equation (8)—and on the ratio θ of discretionary to quantitative skilled investors in the market. This point is key for obtaining the model's main predictions.

4.3 Implications

The analysis yields six main propositions.

4.3.1 *Optimal Learning*

KVV show that, in *recessions*, the marginal benefit of learning about the aggregate shock (λ_n) increases with σ_n and ρ. Whereas in expansions, when σ_n is lower, discretionary investors shift their attention to idiosyncratic shocks, focusing on those with the highest volatility (σ_i).

[10] The solution is obtained through waterfilling. Non-symmetric equilibria where each investors might choose a different attention allocation are possible due to the same substitution effect.

In my model the learning choice of discretionary investors is influenced by the share of skilled funds that are quantitative, θ, and also by how informed those funds are, this is reflected in each shock's marginal benefit of learning (λ_i). Hence, with respect to the KVV result, in my model the incentive to learn about the aggregate shock in recessions is magnified by increases in θ, through two channels: $\frac{\partial \lambda_n}{\partial \sigma_n}$ increases with θ, whereas $\frac{\partial \lambda_i}{\partial \sigma_i}$ $\forall\ i \neq n$ decreases with θ. Additionally, also the incentive to learn about idiosyncratic shocks in expansions decreases with increases in θ ($\frac{\partial \lambda_i}{\partial \sigma_i}$ $\forall\ i \neq n$ decreases with θ); this effect is stronger for assets about which quantitative investors are more informed (with the smallest lower-bound in quant private signal precision $\underline{\sigma}_{iq}^{-1}$).

Proposition 1. *The incentive of discretionary investors to learn about the aggregate shock (λ_n) weakly increases with the share θ of skilled investors who are quantitative. Their incentive to learn about idiosyncratic shocks (λ_i $i \neq n$) weakly decreases with θ—especially for shocks about which precision of the quantitative private signal $\left(\underline{\sigma}_{iq}^{-1}\right)$ is high. For large enough θ, these investors specialize in learning about the aggregate shock.*

4.3.2 *Holdings*

Investors optimally hold more of what they know better. This is evident from equation (6), which shows that the optimal portfolio allocation $\tilde{q}_j^*$ is directly proportional to posterior precision $\hat{\Sigma}_j^{-1}$.

Quantitative funds learn about more shocks, hence have a more precise signal about more of the risky assets; this leads them to hold more of them, as evident from equation (6).

Proposition 2. *Quantitative investors optimally hold a greater number of stocks than do discretionary investors.*

Next I define the information gap (G_i) as the difference in private signal precision, with respect to shock i, between discretionary and quantitative skilled investors:

$$\begin{aligned} G_i &\equiv (K_{iq} - K_{id}) \quad \forall i = 1, ..., n-1, \\ K_{iq} &= \underline{\sigma}_{iq}^{-1} \quad \forall i = 1, ..., n-1. \end{aligned} \tag{13}$$

Note that G_i is generally positive, since quantitative investors have the capacity for learning about a much greater amount of information. It could also be negative, as when the private signal precision of quantitative investors is extremely low because available information is "soft" (i.e., not machine processable).

The marginal benefit of learning about shocks with a greater information gap is more sensitive to the share θ of skilled investors that are quantitative because there is overall more information about that shock in the market ($\bar{K}_i$ increases more). As a result, discretionary investors optimally shift their attention to stocks for which the information gap is smallest.

Proposition 3. *An increase in the share θ of skilled investors that are quantitative weakly reduces the attention allocation of discretionary investors to risk factors with a relatively greater information gap.*

4.3.3 *Dispersion of Opinion*

The dispersion of opinion of a representative quantitative (discretionary) investor with respect to other quantitative (discretionary) investors is given by (derivation in Appendix D.3):

$$E\left[\left(\tilde{q}_q - \bar{\tilde{q}}_q\right)\left(\tilde{q}_q - \bar{\tilde{q}}_q\right)'\right] = \frac{1}{\rho^2}\sum_{i=1}^{n-1} \underline{\sigma}_{iq}^{-1} \tag{14}$$

$$E\left[\left(\tilde{q}_d - \bar{\tilde{q}}_d\right)\left(\tilde{q}_d - \bar{\tilde{q}}_d\right)'\right] = \frac{1}{\rho^2}\sum_{i=1}^{n} \left(\sigma_{id}^{-1} - K_{id}\right)^2 f\left(\sigma_i, \bar{K}_i, \bar{x}_i, \rho\right) + \frac{1}{\rho^2}K \tag{15}$$

where $\bar{\tilde{q}}_q = \int \left[\frac{1}{\rho}\hat{\Sigma}_q^{-1}\left(E_q[\tilde{f}] - \tilde{p}r\right)\partial q\right]$, $\bar{\tilde{q}}_d = \int \left[\frac{1}{\rho}\hat{\Sigma}_d^{-1}\left(E_d[\tilde{f}] - \tilde{p}r\right)\partial d\right]$ and $f\left(\sigma_i, \bar{K}_i, \bar{x}_i, \rho\right)$ is function of the prior volatility of shock i (σ_i), it's expected supply ($\bar{x}_i$), the average private signal precision of all investors in the market about shock i ($\bar{K}_i$) and risk aversion (ρ).[11]

Dispersion of opinion is determined by two effects. First, dispersion of opinion increases with the total precision of private signals: the greater the total precision of private signals the more weight is given to the heterogeneous private signals as opposed to common priors in determining posteriors. Second, dispersion of opinion increases with the cumulative difference in the attention allocated by investors to each asset with respect to the attention allocated to the same assets by the average investor of their type. Risk tolerance magnifies both effects.

For quantitative investors (eq. 14) dispersion of opinion is entirely determined by the first effect; in fact attention allocation is always symmetric among quantitative funds as they optimally learn all available and machine processable information. For what regards discretionary investors (eq. 15), instead, dispersion of opinion is determined by both effects. The first effect is represented by the second term of equation (15): dispersion of opinion is an increasing function of total private signal precision K. The first term of equation (15), instead, represents the second effect: the greater the difference in signal precision of the investor with respect to other discretionary investors $\left(\sigma_{id}^{-1} - K_{id}\right)$, the greater the dispersion of opinion. Discretionary investors, having a limited learning capacity, need to choose what to learn about, due to a substitution effect they might choose to learn about different shocks, this increases dispersion of opinions.

Proposition 4. *As long as $\sum_{i=1}^{n}\left(\sigma_{id}^{-1} - K_{id}\right)^2 f\left(\sigma_i, \bar{K}_i, \bar{x}_i, \rho\right) > \sum_{i=1}^{n-1}\underline{\sigma}_{iq}^{-1} - K$, dispersion of opinion is greater among discretionary investors than among quantitative investors.*

[11] The exact specification can be found in the Appendix

4.3.4 *Performance*

An investor's risk-adjusted performance is measured as the expected excess return with respect to the market return. Appendix E establishes that the excess return of investor j is given by

$$E\left[(R_j - R_M)\right] = E\left[\left(\tilde{q}_j^* - \bar{\tilde{q}}\right)'\left(\tilde{f} - \tilde{p}r\right)\right] = \frac{1}{\rho}\sum_{i=1}^{n}\left[\lambda_i\left(K_{ij} - \bar{K}_i\right)\right]. \tag{16}$$

Expected excess returns increase with the learning advantage of investor j on the risk factors it chooses to learn about $\left(K_{ij} - \bar{K}_i\right) > 0$—and in proportion to the marginal benefit λ_i of learning about them.

With regard to the aggregate shock, quantitative investors are always at a learning disadvantage compared with discretionary investors. With regard to idiosyncratic shocks, however, quantitative investors generally have a learning advantage over all other investors in the market—provided G_i is positive. When G_i is negative, their advantage is reduced but it remains positive as long as the total share of skilled investors in the market is not too high $\left(\frac{\underline{\sigma}_{iq}^{-1}}{K_{id}} > \frac{\chi-\chi\theta}{1-\chi\theta}\right)$:

$$\begin{cases} \left(\underline{\sigma}_{iq}^{-1} - \bar{K}_i\right) = \left[(1-\chi)K_{id} + (1-\chi\theta)G_i\right] > 0, & i = 1, ..., n-1, \\ \left(\sigma_{nq}^{-1} - \bar{K}_i\right) = -\left[K_{id}(\chi - \chi\theta)\right] < 0, & i = n. \end{cases} \tag{17}$$

The performance of quantitative investors increases in tandem with increases in σ_i for $i = 1, ..., n-1$ thanks to the increase in the marginal benefit of learning about those risk factors for which they have a learning advantage. Their excess return is greatest when discretionary skilled investors choose *not* to learn about the aggregate risk factor. If instead they do choose to learn about it, then the excess return of quantitative investors declines following increases in σ_n. Finally, their performance deteriorates as θ increases. The negative effect of an increase in θ on excess returns is greatest when the elasticity of λ_i with respect to θ is high. Elasticity is increasing in $\sum_{i=1}^{n-1}\underline{\sigma}_{iq}^{-1}$, from which it follows that excess returns deteriorate faster when each fund has higher precision of private signals.

Proposition 5. *The excess returns of quantitative skilled investors increases with σ_i for $i = 1, ..., n-1$, weakly decreases in recessions (when σ_n increases), and always decreases with the share θ of quantitative investors.*

Greater signal precision $\underline{\sigma}_{iq}^{-1}$ acts in two opposite directions. It directly increases excess returns thanks to the fund's greater informativeness in choosing its portfolio allocation. It indirectly decreases excess returns by reducing the marginal benefit of learning about the same shocks, through the elasticity of λ_i with respect to θ. This is particularly relevant as quantitative investors are unconstrained in their learning capacity, hence even small increases in the share of quantitative investors θ can have a great impact on λ_i.

For what regards discretionary investors, when choosing to learn about the aggregate risk factor, they earn a positive excess return if their informational advantage with respect to both unskilled investors and quantitative investors (i.e., their ability to learn about the aggregate shock—the "flexibility advantage" captured by the LHS of equation 18) is *greater* than their learning disadvantage with respect to quantitative investors (i.e., owing to their smaller information capacity—the "breadth disadvantage" captured by the RHS of equation 18). Formally,

$$\lambda_n \left[K - (\chi(1-\theta))K_{nd}\right] > \chi\theta \sum_{i=1}^{n-1} \left[\lambda_i \underline{\sigma}_{iq}^{-1}\right]. \tag{18}$$

When, instead, discretionary investors optimally choose to learn about idiosyncratic shocks, their expected excess return is positive if the following (more restrictive) condition is verified:

$$\lambda_l \left(K - \chi K_{ld} - \chi\theta G_l\right) > \chi\theta \sum_{i \neq j}^{n-1} \left\{\lambda_i \left[\underline{\sigma}_{iq}^{-1}\right]\right\}. \tag{19}$$

In this latter case, they have a learning advantage only with respect to unskilled investors; in contrast, they display a learning disadvantage, which is equal to the information gap G_l, with respect to quantitative investors. Even when the information gap is negative, equation (19) represents a more restrictive condition, since the aggregate shock is in much greater supply than any idiosyncratic shock l that they might choose to learn about. Furthermore, an increase in the prior volatility of the idiosyncratic shocks l, which they choose to learn about, increases their expected excess return only if θ is sufficiently low:

$$\theta < \frac{\sigma_{ld}^{-1} - \chi K_{ld}}{\chi G_l}. \tag{20}$$

Note that, when discretionary investors choose to learn about shocks with a smaller information gap, their excess return rises for a wider range of θ values – potentially being always verified if focusing on shocks with . Equations (19) and (20) highlight the direct link between performance and information gap reduction. Additionally, according to equation (20), there exist values of $\theta < 1$ for which discretionary investors *cannot* earn positive excess returns from learning about idiosyncratic shocks. This is what leads to specialization in the aggregate shock, per Proposition 1, for high enough values of θ.

When discretionary investors optimally choose to learn about the aggregate shock, increases in σ_n always have a positive effect on their excess return whereas increases in σ_i, $i = 1, \ldots, n-1$, (marginally) reduces it. This observation, together with equations (18) and (19), leads us to conclude that discretionary skilled investors always have higher expected excess returns in recessions.

Finally, when discretionary investors choose to learn about idiosyncratic risks, their excess return falls as θ rises. If instead they choose to learn about the aggregate risk, then increases in θ lead to an increase in their expected excess return—provided the elasticity of λ_i $(i-1, \ldots, n)$ with respect to θ is sufficiently large–this condition is not satisfied for shocks with a negative information gap.

Proposition 6. *When discretionary investors, optimally learn about idiosyncratic shocks, their excess return decreases with θ and with the volatility of the idiosyncratic risk factors that they ignore.*

The excess return increases with the volatility of the shocks on which they focus but only if θ and G_l are both low enough. When discretionary investors optimally learn about the aggregate shock, their excess return always increases with σ_n and decreases with $\sigma_{i\neq n}$; it also increases with θ if the elasticity of $\lambda_{i\neq n}$ and of λ_n (with respect to θ) are sufficiently high.

4.3.5 *Price Informativeness*

We can use equation (8) to rewrite the average precision of private signals about any idiosyncratic shock i, in terms of G_i, as follows:

$$\bar{K_i} = \chi K_{id} + \chi\theta G_i. \tag{21}$$

This formulation reflects that the average private signal precision about idiosyncratic shocks amounts to the sum of the discretionary signal's precision, χK_{id}, and a premium ($G_i > 0$) for the additional signal precision of quantitative funds, $\chi\theta G_i$. However, the presence of quantitative investors may end up reducing the average private signal precision about stocks for which there is little machine-learnable information ($G_i < 0$).

I define *price informativeness* as the precision of equilibrium prices, which is given by

$$\left(\Sigma_p^{-1}\right)_{ii} = \sigma_{pi}^{-1} = \frac{\bar{K_i}^2}{\rho^2\sigma_x} \quad \forall i = 1, ...n. \tag{22}$$

This expression is an increasing function of the average private signal precision $\bar{K_i}$ of skilled investors, so it always increases with χ (the share of skilled investors in the market). However, the effect of increases in the share θ of quantitative investors depends on the type of shock, on the extent of the information gap, and on the state of the business cycle.

Proposition 7. *An increase in θ increases (resp. decreases) the price informativeness of idiosyncratic shocks with a high (resp. low) information gap. An increase in θ decreases (resp., weakly increases) the price informativeness of the aggregate shock in recessions (resp., in expansions).*

The first part of Proposition 7 follows from equations (21) and (22) when one considers the effect of θ on the learning choice of discretionary investors. For the second part, the intuition is as follows. In recessions, when all skilled investors are discretionary, they all learn about the aggregate shock; but if some skilled investors are quantitative, fewer skilled investors learn about that shock, reducing its price informativeness. In expansions, increases in θ have no effect on the aggregate shock's price informativeness provided that discretionary funds choose to learn about idiosyncratic shocks. Yet for large enough θ they optimally choose to specialize in learning about the aggregate shock (Proposition 1), in which case its price informativeness is increased.

5 Empirical Analysis

5.1 Empirical Measures

The stock-picking and market-timing measures are constructed as in KVV. Stock picking is measured as the covariance between the weight allocated to each stock by the fund (in excess of the market weight) and the next period's stock-specific shocks. Market timing is measured as the 12-month rolling covariance between the weight allocated to each stock by the fund (in excess of the market weight) and the next period's aggregate shock. Formally:

$$Timing_{jt} = \frac{1}{TN^j} \sum_{i=1}^{N^j} \sum_{\tau=0}^{T-1} (w^j_{i,t+\tau} - w^m_{i,t+\tau})(b_i z_{n,(t+\tau+1)}), \tag{23}$$

$$Picking_{jt} = \frac{1}{N^j} \sum_{i=1}^{N^j} (w^j_{i,t} - w^m_{i,t})(z_{i,(t+1)}). \tag{24}$$

Here j denotes the fund, i the stock, and N^j the number of stocks held by fund j; w^j_{it} is the weight (measured as a percentage of TNA) of stock i in the portfolio of fund j at time t, and w^m_{it} is the weight (measured as a percentage of total market capitalization) of stock i at time t in the CRSP universe. The term z_{it} (resp., z_{nt}) is the shock specific to stock i (resp., the aggregate shock) at time t, and b_i is the exposure of stock i to the aggregate shock.[12] The principle behind these measures is that a fund's holdings reflect the information learned by its manager. Hence, for skilled managers, holdings should covary with the future fundamentals of either stock-specific shocks or the aggregate shock—depending on the particular shocks about which the manager has private information. How well the manager chooses portfolio weights in anticipation of future asset-specific or aggregate shocks is closely related to (respectively) her stock-picking and market-timing abilities. I use earnings "surprises" (SUE) to measure stock-specific shocks and growth in industrial production to measure the aggregate shock as in KVV.

The information gap is measured as the average availability of machine-processable information for the stocks held by each fund over time. I consider three types of averages: simple average; allocation-weighted average, where the weights are the percentage of TNA allocated to stocks; and allocation difference–weighted average, where the weights are the normalized weights of each stock in the fund's portfolio (in excess of its market weight). I refer to this last measure as the *active information gap*:

$$InfoGap^{wd}_{jt} = \sum_{i=1}^{N^j} \left[\frac{|w^j_{i,t}\text{-}\ w^m_{i,t}|}{\sum_{i=1}^{N^j} |w^j_{i,t}\text{-}w^m_{i,t}|} I_{it} \right] \tag{25}$$

Here $w^j_{i,t} - w^m_{i,t}$ is the weight (in excess of its market weight) of stock i in the portfolio of fund j

[12] b_i is obtained through a 12-months rolling regression of the return of stock i on future aggregate shocks.

at time t, and I_{it} is a measure of machine-processable information availability for stock i at time t. I expect the learning effect to be best captured by this active information gap, since measuring information availability for stocks that the manager decides to underweight or overweight (with respect to the market) should better reflect her active learning.

I construct different measures of the information gap using different proxies for the amount of information available about stocks (I_{it}). I use the stock's size (market capitalization) and age (in months), number of IBES end-of-period forecasts for the stock, and number of times the stock was mentioned in Dow Jones news articles per month. The logic underlying these proxy choices is that there should be less information available—and machine processable—about smaller and younger stocks, which are less followed by analysts and in the media.

I measure the share θ of quantitative investors in two ways (i) as the share of assets managed by quantitative funds relative to the total assets managed by funds in my sample (θ^{TNA}); and (ii) as the percentage of the US stock market capitalization held by quantitative funds (θ^{Hold}):

$$\theta_t^{TNA} = \frac{TNA_t^{Quantitative}}{TNA_t^{Quantitative} + TNA_t^{Discretionary}} \tag{26}$$

$$\theta_t^{Hold} = \frac{\sum_{i=1}^{Q_t} \sum_{i=1}^{N_t^j} |n_{ijt}| P_{it}}{\sum_{i=1}^{N^j} n_{it} P_{it}} \tag{27}$$

Here j denotes the fund, i the stock, Q_t the total number of quantitative funds at time t and N_t^j the number of stocks held by fund j at time t; P_{it} is the price of stock i at time t, n_{ijt} is the number of shares of stock i held by fund j at time t and n_{it} is the number of shares outstanding of stock i at time t.

Finally, I also build two measures of dispersion of opinion. First I measure portfolio overlap (i.e. commonality) for funds of the same group–quantitative or discretionary–as the allocation-weighted average of the percentage of funds of the same group which hold each stock over time. For the same reasons explained in the construction of the Active Information Gap measure, weights are the normalized percentages of the fund's TNA allocated in each stock $w_{it}^j - w_{it}^m$ (in excess of their market weight):

$$Comonality_{jt}^F = \sum_{i=1}^{N_t^j} \left[\frac{|w_{it}^j - w_{it}^m|}{\sum_{i=1}^{N_t^j} |w_{it}^j - w_{it}^m|} \left(\frac{F_{it}}{F_t} \right) \right]; \tag{28}$$

for $F = (Q,\ D)$, where Q_t (D_t) is the total number of quantitative (discretionary) funds at time t and Q_{it} (D_{it}) is the number of quantitative (discretionary) funds who hold stock i at time t.

Second I measure holdings dispersion as the cumulative squared difference in the weight allocated by each funds and the average weight allocated by funds of the same type to stocks.

$$Dispersion_{jt}^{F} = \sum_{i=1}^{N_t^j} \left(w_{it}^{j} - \bar{w}_{it}^{F} \right)^2$$

for $F = (Q,\ D)$, where $\bar{w}_{it}^{F}$ is the average weight allocated by quantitative (discretionary) funds to stock i at time t.

What I subsequently call “overcrowding” is a combination of the prevalence of quantitative funds in the market (θ) and the commonality in their portfolios

5.2 Tests of the Model Assumptions

As a first step, I provide evidence in support of the model's main assumptions: that skilled *quantitative* funds specialize in learning about idiosyncratic shocks; and that the funds I classify as *discretionary* adjust what they learn about, as KVV find.

In KVV, skilled investors are limited in information capacity but perfectly flexible in shifting their attention from aggregate to idiosyncratic shocks. The authors find that skilled investors switch from stock picking in expansions to market timing in recessions. I test for whether the skilled funds in my sample display this switching behavior and for whether the effect differs between quantitative and discretionary funds. For that purpose, I identify funds that are most persistently among the top stock pickers in expansions and then check to see whether they switch to market timing in recessions. Finally, I check for whether any of the fund classes specializes in either stock picking or market timing *across* phases of the business cycle.

I begin by constructing the market-timing and stock-picking measures described in Section (5.1). Next I identify all observations belonging to the top quantile of the market-timing distribution in recessions or of the stock-picking distribution in expansions (i.e., the high-ability sample). Finally, for each fund I calculate how frequently members of this high-ability sample exhibit market-timing ability in recessions or stock-picking ability in expansions. Frequency is measured as the number of months (relative to their respective lifetimes) that funds are among the top timers or the top pickers. Thus $TopPickers^{E}$ and $TopTimers^{R}$ refer to the top quantile of funds that most frequently display, respectively, high stock-picking ability in expansions and high market-timing ability in recessions.

I then run the following regressions, where (as before) $Quant_j$ is a dummy set equal to 1 if the fund is quantitative (and to 0 otherwise):

$$Timing_{jt} = \alpha + \beta_1 TopPickers_j^{E} + \beta_2 Quant_j + \beta_3 TopPickers_j^{E} * Quant_j + \gamma X_{jt} + \epsilon_t | Recession; \quad (29)$$

$$Picking_{jt} = \alpha + \beta_7 TopPickers_j^{E} + \beta_8 Quant_j + \beta_9 TopPickers_j^{E} * Quant_j + \gamma X_{jt} + \epsilon_t | Recession. \quad (30)$$

$$Timing_{jt} = \alpha + \beta_4 TopTimers_j^R + \beta_5 Quant_j + \beta_6 TopTimers_j^R * Quant_j + \gamma X_{jt} + \epsilon_t | Expansion; \quad (31)$$

Recessions and expansions are identified using NBER business cycle dates. The term X_{jt} represents a series of fund-specific control variables previously described (see Table 1 for details).

Equation (29) tests whether high-ability funds switch from stock picking in expansions to market timing in recessions. Consistently with the KVV results, I find that the funds with high stock-picking ability in expansions also have high market-timing ability in recessions ($\beta_1 = 0.454 > 0$). For quantitative funds, however, this effect cancels ($\beta_3 = -0.621 < -\beta_1$): quantitative funds do not optimally switch between stock picking and market timing across the business cycle (Model 1 in Table 5).

Equations (30) and (31) test for whether some funds specialize in, respectively, stock picking and market timing across the business cycle. In line with the model's assumptions, I find that quantitative funds among the top stock pickers in expansions are also among top stock pickers in recessions ($\beta_8 = 0.356 > 0$; $\beta_9 = 0.248 > 0$); these results confirm that quantitative funds specialize in learning about idiosyncratic shocks (Model 2 in Table 5). Finally no funds are persistently among the top market timers across business cycle phases (i.e., neither β_4 nor β_6 is significant). This finding confirms the hypothesis that, among equity mutual funds, none specializes in market timing (Model 3 in Table 5).

5.3 Market-Timing and Stock-Picking

Turning to the model's main predictions, I first test Proposition 1: as the share of quantitative investors rises, the incentive of discretionary investors to learn about the aggregate shock in recessions increases; conversely, their incentive to learn about idiosyncratic shocks in expansions decreases. I test these predictions via the measures (described in Section 5.1) of market timing, stock picking, and share of quantitative funds. Particularly the share of quantitative funds is measured as a proportion of assets managed (θ^{TNA}) and as the percentage of the US market capitalization held by quantitative funds (θ^{Hold}). These two measures are used in running the following regressions:

$$\begin{aligned} Picking_{jt} &= \alpha + \beta_1 NBER_t + \beta_2 \theta_t^{Type} + \gamma X_{jt} + \epsilon_t | \quad j = d \\ Timing_{jt} &= \alpha + \beta_3 NBER_t + \beta_4 \theta_t^{Type} + \gamma X_{jt} + \epsilon_t | \quad j = d \end{aligned}$$

where $NBER_t$ is an indicator variable which identifies recession periods, θ^{Type} (for $Type = (TNA,\ Hold)$) is the proxy for θ and X_j are a series of fund specific control variables (The same as specified in table 1).

I first confirm KVV's finding that the stock-picking ability of discretionary investors is greatest in expansions ($\beta_1 < 0$) whereas their market-timing ability is greatest in recessions ($\beta_3 > 0$).

I also find evidence in favor of Proposition 1. Using θ^{TNA} as a proxy for θ, the coefficient $\beta_2 = -0.143$ is negative and significant at the 1% level: a 1% increase in θ reduces the stock-picking ability of discretionary investors by 8.5% of a standard deviation. Conversely, $\beta_4 = 0.149$ is positive and significant at the 1% level: a 1% increase in θ increases the market timing ability of discretionary investors by 14% of a standard deviation (Models 1–2 of Table 8).

A concern is reverse causality if θ^{TNA} proxies for outflows from discretionary funds. For that to be the case, we would first need to believe in a direct causality between outflows and learning ability, additionally for the signs to be correct it would imply that outflows impact stock picking positively and market timing negatively, this seems implausible. To further alleviate this concern, I repeat the analysis using θ^{Hold} (eq. 27), which, despite not being completely independent from fund flows, is not directly related to discretionary fund flows. In fact funds might change their investment in the stock market to hold more cash or invest in other non-equity products. I obtain similar findings (Table 8, Models 3–4). Results are robust controlling for the total US stock market capitalization held by discretionary funds.

5.4 Portfolio Diversification

Next I test Proposition 2, which states that quantitative funds are expected to hold significantly more stocks than do discretionary funds.

As reported in Table 6, quantitative funds hold about double the number of stocks held by discretionary funds ($\simeq$117+108 vs. 117). Results are robust to controlling for the amount of cash held, the size of funds, and various other fund characteristics. Moreover, the distribution of the number of stocks held by discretionary funds is quite concentrated, with most observations falling between 10 and 100 stocks. Quantitative funds exhibit a more dispersed distribution, and a sizable number of them hold as many as 1,000 stocks. The distribution of the number of holdings seems not to vary with recessions and expansions (Figure 10). [13]

The greater portfolio diversification among quantitative funds is reflected in their marginally lower portfolio risk, as shown in Table 7. The first two columns consider the 12- and 36-month rolling volatility of fund returns. Compared with discretionary funds, quantitative funds exhibit (on average) 2% lower return volatility; and in recessions, volatility increases 8% less. In columns 3–8 of the table, portfolio risk is defined as the 12- and 36-month rolling *idiosyncratic* volatility of fund returns. Idiosyncratic volatility is measured as the volatility of the residual of a regression of excess fund returns on (i) the excess return on the market, (ii) the Fama–French three-factor model, and (iii) the Carhart four-factor model (CAR). I find that quantitative funds' returns display between 8.3% (for 36-month rolling CAPM) and 15.6% (for 12-month rolling CAR) less idiosyncratic volatility than discretionary funds.

[13] Quantitative funds hold a significantly larger number of stocks than discretionary funds also when using quantile regressions based on the distribution of the number of holdings – results available upon request

5.5 Information Gap

Proposition 3 indicates the presence of cross-sectional differences in the learning of discretionary and quantitative investors; in particular, skilled discretionary investors should shift their attention to stocks for which less information is learnable/available in order to reduce their information gap with respect to quantitative investors. For each proxy I construct the simple average, the allocation-weighted average, and the active information gap (equation 25) for the stocks held by each fund over time. I then regress the information gap proxies on dummies that identify recession periods and discretionary funds, an interaction between those two dummies, and various fund-specific characteristics:

$$InfoGap_{jt} = \alpha + \beta_1 NBER_t + \beta_2 Discretionary_j + \beta_3 NBER_t \times Quant_j + \gamma X_{jt} + \epsilon_t;$$

here $NBER_t$ is a dummy variable identifying recessions, $Discretionary_j$ is a dummy variable that identifies discretionary funds, and X_{jt} represents the fund-specific control variables described in Table 1.

I expect β_2 to be negative and significant, which would indicate that (on average) discretionary funds hold stocks with a smaller information gap. Mutual fund returns are evaluated against a benchmark, so one should expect mutual fund managers to hold not only stocks about which they actively learn but also a combination of other stocks that replicate the benchmark. In that case, active mutual fund managers can be expected to underweight or overweight—with respect to that benchmark—those stocks for which they have (through learning) formed a directional view. That dynamic could bias upward the simple average measure of information gap when stocks held as part of the benchmark are large ones for which a substantial amount of information is available. This measure should have an even greater upward bias for discretionary funds given that, as shown in Section 5.4, they hold fewer stocks. For that reason I expect the simple average measure of the information gap to fare the worst at capturing this effect. The allocation-weighted average should mitigate such an upward bias, but the best measure should be the active information gap because it isolates those stocks for which the manager's portfolio allocation deviates from the benchmark, an indication of active learning.

I find that discretionary funds hold 17% smaller stocks (Model 3 in Table 9). As expected, the effect is not significant for the simple average measure of the information gap (Model 1) but is significant when either the allocation-weighted average or the active information gap (Models 2 and 3, respectively) is used. In addition, discretionary funds hold stocks that are 11% younger (Model 5 in Table 10); this effect persists even after controlling for the average stock size, which indicates that stock age is not just a proxy for stock size. The age effect is significant for all measures, but it is stronger when the allocation-weighted average or the active information gap are used. Discretionary funds also hold stocks that have 11% fewer mentions in Dow Jones media (Model 5 of Table 11). This effect is not significant when I use the simple average but is significant under the allocation-weighted average—though only at the 10% level when controlling for average market capitalization of the stocks held. As expected, the effect is strongest (11% fewer mentions; significant at the 1% level) under the active information gap measure, even after controlling for the average market capitalization of the stocks held. The effect of discretionary funds on the information gap, as measured by the number of analyst forecasts, is negative but not significant (Model 5 in Table 12).

Finally, we confirm that discretionary funds earn a marginally higher premium when investing in stocks characterized by a smaller information gap (Table 13).

5.6 Dispersion of Opinion

Proposition 4 states that there should be a greater dispersion of opinion among discretionary funds than among quantitative funds. I test this proposition using the measures for commonality and dispersion described in section 5.1. The intuition being that a greater dispersion of opinion among investors about assets' payoff should lead to more diversified hidings. Figure 11 displays the distribution of commonality (Panel 1) and holdings dispersion (Panel 2) among quantitative and discretionary funds over the entire period. Quantitative funds display a lower dispersion (more concentrated distribution) and higher holdings commonality than discretionary funds. Additionally we can clearly see a two-picked distribution for the commonality measure, particularly pronounced for quantitative funds, indicating that funds polarize between a low and high common portfolio ownership, which for quantitative funds picks at roughly 6% and 20%. Table 14 further shows the statistical significance of this difference. Model 1 shows that dispersion of holdings is significantly about 30% greater for discretionary funds than quantitative funds. Whereas Model 2 shows that holdings commonality is significantly lower for discretionary funds (9% vs 14.79%).

5.7 Performance

Propositions 5 and 6 concern the relative performance of quantitative and discretionary funds across the business cycle. In KVV the authors report a significantly higher performance in recessions—34 basis points when measured by 12-month rolling CAPM alphas—and the premium was greatest for funds that shifted from market timing in recessions to stock picking in expansions (their period of analysis was 1980–2005). When replicating that result using my data sample (for the period 1999–2015), I find a lower recession premium (28 bp) if I do not control for the $Quant_j$ dummy (Model 1 in Table 15). Yet when controlling for the presence of quantitative funds, the recession premium of discretionary investors is recovered (31.4 bp), and it emerges that quantitative investors display significantly lower performance: earning 15.7 bp less than discretionary investors in recessions (Model 2 in Table 15). Results are robust to incorporating various fund-level controls: age, size, turnover rate, expenses ratio, net flows, fund flow volatility, fund loads, and exposure to the size, value, and momentum factors (Model 3 in Table 15). Results persist also when computing performance as the 36-month rolling CAPM alpha, despite the recession's effect being weaker under that criterion (Model 4 in Table 15).

Similar results are obtained when performance is computed using the Fama–French three-factor model (Table 16) and the Carhart four-factor model (Table 17); these results are in line with Propositions 5 and 6. With respect to the Fama–French three-factor 12-month rolling alphas, the performance premium is not significant absent controlling for the different performance of quantitative and discretionary funds. A recession premium of 8.56 bp is recovered for discretionary funds when the regression includes the $Quant_j$ dummy, in which case quantitative funds are found to return (on average) 9.51 bp less than discretionary investors in recessions. Results are robust to adding fund-specific control variables and using 36-month rolling alphas. When using the Carhart four-factor model, we observe a recession premium of 4.26 bp that is significant only when using 36-month rolling alphas; in contrast, the negative premium of quantitative funds in recession—which ranges from 6.04 bp to 9.01 bp—is significant under all model specifications.

Regarding the performance of quantitative funds, Proposition 5 highlights different effects that go in opposite directions. First, the learning behavior of quantitative funds induces them to make more informed trades in expansions. Second, the greater amount of information learned, leads to greater portfolio diversification, as shown in Section 5.4. This should produce better risk-adjusted performance, but at the same time it should prevent them from being greatly exposed to any single idiosyncratic shock from which they could profit. Finally performance deteriorates more the greater is the share of quantitative investors in the market and information availability ("overcrowding"). Overcrowding is exacerbated when funds hold similar portfolios.

In order to disentangle these different effects, I have looked at the two dimensions of "overcrowding": holdings commonality and the rise in the share of quantitative funds over time. Particularly the greater overlap in the quantitative portfolios holding leads "overcrowding" to have a stronger impact on the performance of quantitative funds (Figure 11). In order to further explore the effect of overcrowding over time I have built two additional dummy variables: E_1 and E_2 which identify respectively the expansion period prior to the millennial recession (Dec 1999–Mar 2001) and the expansion period in between the two recessions in my sample (Dec 2001–Nov 2007) . The regression specifications also contain a recessions dummy, hence all coefficients are interpreted relative to the expansion post subprime crisis:

$$\alpha_{jt}^{Q} = \gamma + \gamma_1 E_{1t} + \gamma_2 E_{2t} + \gamma_3 NBER_t + \epsilon_t \tag{32}$$

where α_{jt} is the 12-months rolling CAPM, Fama-French 3-factor and Carhart 4-factor alpha.

Models 1,2 and 3 of Table 18 clearly show that the expansion premium of quantitative funds has been decreasing over time. In the CAPM specification, the per-millennial recession (E_1) expansion displays a premium of 79.8bp more than the post-subprime expansion, the premium fuhrer decrease to 12.1bp for the per-subprime second expansion. In the CAPM specification a recession premium of 20.7bp also figures but this becomes insignificant when using the 3-factor or 4-factor models. For the 3-factors and 4-factors models the E_1 premium is of 40.4bp and 46.0bp respectively, while the E_2 expansion doesn't seem to be significantly different from the post-subprime one. Model 4 of Table

18 runs the same specification as in equation (32) but using the 12-months rolling Sharpe ratio as independent variable. By doing so I find a significantly positive Sharpe Ratio in expansions (25.1bp), while a negative Sharpe ratio emerges for E_1 (58.1bp–significant at the 10% level). Additionally recessions display a significantly negative Sharpe ratio (53.2bp). This suggests that quantitative funds might have countered the worsening in performance due to overcrowding with progressively better risk management practices. Finally Models 5 and 6 of Table 18 show that quantitative funds with a smaller commonality in holdings tend to perform better. The result is strong both using the commonality measure (Model 5) and a commonality dummy, which takes a value of 1 to identify those funds who display commonality above the median for their group (Model 6).

6 Conclusion

This paper demonstrates that the distinction between quantitative and discretionary funds is extremely relevant. The key to understanding what distinguishes them lies in their respective learning behaviors and in the relative share of quantitative to discretionary funds in the market. The current status quo in quantitative modeling determines a trade-off between information quantity and flexibility, which allows human traders to carve a niche in which they can maintain significantly greater (risk-adjusted) performance in recessions. Quantitative investors, though, display better portfolio diversification and risk management throughout the business cycle. A very important concept in understanding the relative performance of quantitative and discretionary funds is overcrowding. As the amount of information progressively incorporated in prices increases, the ability of investors to maintain high risk adjusted performance deteriorates. This is particularly relevant as quantitative investors are unconstrained in their learning capacity, hence even small increases in the share of quantitative investors in the market can have a great impact on performance. Funds who are able to achieve a lower overlap of their portfolio with respect to their peers are able to maintain a competitive edge. It is important to understand this interaction because it could foretell future developments. Indeed, increases in the availability and use of big data could well accentuate the breadth advantages of quantitative funds but it could also worsen the overcrowding effect, unless funds are able to maintain heterogeneity in their portfolio allocations. At the same time, recent developments in artificial intelligence—in particular, deep learning neural networks—could provide a foundation for overcoming the limited flexibility of current quantitative models. This paper provides a framework to start thinking about these issues.

References

[1] Anat R. Admati. A noisy rational expectations equilibrium for multiasset securities markets. *Econometrica*, 53:629–657, 1985.

[2] Parvez Ahmed and Sudhir Nanda. Performance of enhanced index and quantitative equity funds. *Financial Review*, 40(4):459–479, 2005.

[3] Miguel Anton and Christopher Polk. Connected stocks. *The Journal of Finance*, 69(3):1099–1127, 2014.

[4] Casey Quirk & Associates. The geeks shall inherit the earth? quantitative managers' recent edge. *Research Insight*, 2005.

[5] Malcolm Baker, Lubomir Litov, Jessica A. Wachter, and Jeffrey Wurgler. Can mutual fund managers pick stocks? evidence from their trades prior to earnings announcements. *Journal of Financial and Quantitative Analysis*, 45(5):1111–1131, 2010.

[6] Nataliya Bershova and Dmitry Rakhlin. High-frequency trading and long-term investors: a view from the buy-side. *Journal of Investment Strategies*, 2(2):25–69., 2013.

[7] Ekkehart Boehmer, Kingsley Y. L. Fong, and Juan Wu. International evidence on algorithmic trading. *Available at SSRN*, 2015.

[8] Jonathan Brogaard, Terrence Hendershott, and Ryan Riordan. High-frequency trading and price discovery. *Review of Financial Studies*, 27(8):2267–2306, 2014.

[9] John Y. Campbell, Martin Lettau, Burton G. Malkiel, and Yexiao Xu. Have individual stocks become more volatile? an empirical exploration of idiosyncratic risk. *The Journal of Finance*, 56(1):1–43, 2001.

[10] Allen Carrion. Very fast money: High-frequency trading on the nasdaq. *Journal of Financial Markets*, 16(4):680–711, 2013.

[11] Alain P. Chaboud, Benjamin Chiquoine, Erik Hjalmarsson, and Clara Vega. Rise of the machines: Algorithmic trading in the foreign exchange market. *The Journal of Finance*, 69(5):2045–2084, 2014.

[12] Yeung L. Chan and Leonid Kogan. Catching up with the jonases: Heterogeneous preferences and the dynamics of asset prices. *Journal of Political Economy*, 110(6):1255–1285, 2002.

[13] Vinod Chandrashekaran. *The Evolution of Quantitative Investment Strategies*. Investment Viewpoints. American Century Investments, March 2015.

[14] Michael Chui, James Manyika, and Mehdi Miremadi. *Where machines could replace human and where they can't (yet)*. McKinsey Quarterly, July 2016.

[15] John H. Cochrane. Financial markets and the real economy. In *International Library of Critical Writings in Financial Economics*, volume 18. London: Edward Elgar, 2006.

[16] Jennifer Conrad, Sunil Wahal, and Jin Xiang. High-frequency quoting, trading, and the efficiency of prices. *Journal of Financial Economics*, 16(2):271–291, 2015.

[17] Thomas M. Cover and Thomas A. Joy. *Elements of Information Theory.* John Wiley and Sons, New York, New York,, first edn. edition, 1991.

[18] Martijn K.J. Cremers and Antti Petajisto. How activeis your fundmanager?a new measure that predicts performance. *Review of Financial Studies*, 22(9):3329–3365, 2009.

[19] Andre De Souza and Anthony W. Lynch. Does mutual fund performance vary over the business cycle? *National Bureau of Economic Research*, (w18137), 2012.

[20] Bernard Dumas. Two-person dynamic equilibrium in the capital markets. *Review of Financial Studies*, 2(2):157–188, 1989.

[21] Edwin J. Elton, Martin J. Gruber, and Christopher R. Blake. A first look at the accuracy of the crsp mutual fund database and a comparison of the crsp and morningstar mutual fund databases. *The Journal of Finance*, 56(6):2415–2430, 2001.

[22] R. Engle and J. Rangel. The spline-garch model for low-frequency volatility and its global macroeconomic causes. *Review of Financial Studies*, 21(3):1187–1222, 2008.

[23] Frank J. Fabozzi and Sergio M. Focardi. Black swans and white eagles: on mathematics and finance. *Mathematical Methods of Operations Research*, 69(3):379–394, 2009.

[24] Frank J. Fabozzi, Sergio M. Focardi, and Caroline L. Jonas. Trends in quantitative equity management: survey results. *Quantitative Finance*, 7(2):115–122, 2007.

[25] Frank J. Fabozzi, Sergio M. Focardi, and Caroline L. Jonas. Challenges in quantitative equity management. *Research Foundation of CFA Institute Publications*, pages 1–110, 2008. (corrected July 2008).

[26] Frank J. Fabozzi, Sergio M. Focardi, and Caroline L. Jonas. On the challenges in quantitative equity management. *Quantitative Finance*, 8(7):649–665, 2008.

[27] Eugene F. Fama and Kenneth R. French. Business conditions and expected returns on stocks and bonds. *Journal of Financial Economics*, 25(1):23–49, 1989.

[28] Lily H. Fang, Joel Peress, and Lu Zheng. Does media coverage of stocks affect mutual funds' trading and performance? *Review of Financial Studies*, 27(12):3441–3466, 2014.

[29] World Economic Forum. *The future of financial services.* An Industry Project of the Financial Services Community. Prepared in collaboration with Deloitte. 2015.

[30] Nicolae Garleanu and Stavros Panageas. Young, old, conservative, and bold: The implications of heterogeneity and finite lives for asset pricing. *Journal of Political Economy*, 123(3):670–685, 2015.

[31] Paul Glasserman and Harry Mamaysky. Market efficiency with micro and macro information. *Available at SSRN*, 2016.

[32] Vincent Glode. Why mutual funds underperform. *Journal of Financial Economics*, 99(3):546–559, 2011.

[33] Robin Greenwood and David Thesmar. Stock price fragility. *Journal of Financial Economics*, 102(3):471–490, 2011.

[34] Sanford J. Grossman and Joseph E. Stiglitz. On the impossibility of informationally efficient markets. *The American economic review*, 70(3):393–408, 1980.

[35] James D. Hamilton and Gang Lin. Stock market volatility and the business cycle. *Journal of Applied EconometricsJournal of Applied Econometrics*, 11(5):573–539, 1996.

[36] Joel Hasbrouck and Gideon Saar. Low-latency trading. *Journal of Financial Markets*, 16(4):646–679, 2013.

[37] Jennifer Huang, Clemens Sialm, and Hanjiang Zhang. Risk shifting and mutual fund performance. *Review of Financial Studies*, 24(8):2575–2616, 2011.

[38] Boyan Jovanovic and Albert J. Menkveld. Middlemen in limit order markets. *Available at SSRN*, 2016.

[39] Marcin Kacperczyk and Amit Seru. Fund manager use of public information: New evidence on managerial skills. *The Journal of Finance*, 62(2):485–528, 2007.

[40] Marcin Kacperczyk, Clemens Sialm, and Lu Zheng. On the industry concentration of actively managed equity mutual funds. *The Journal of Finance*, 60(4):1983–2011, 2005.

[41] Marcin Kacperczyk, Clemens Sialm, and Lu Zheng. Unobserved actions of mutual funds. *Review of Financial Studies*, 21(6):2379–2416, 2008.

[42] Marcin Kacperczyk, Stijn Van Nieuwerburgh, and Veldkamp Laura. A rational theory of mutual funds attention allocation. *Econometrica*, 84(2):571–626, 2016.

[43] Marcin Kacperczyk, Stijn Van Nieuwerburgh, and Laura Veldkamp. Time varying fund manager skill. *The Journal of Finance*, 69(4):1455–1484, 2014.

[44] Amir E. Khandani and Andrew W. Lo. What happened to the quants in august 2007? evidence from factors and transactions data. *Journal of Financial Markets*, 14(1):1–46, 2011.

[45] Ralph S. J. Koijen. The cross-section of managerial ability, incentives, and risk preferences. *The Journal of Finance*, 69(3):1051–1098, 2014.

[46] Robert Kosowski. Do mutual funds perform when it matters most to investors? us mutual fund performance and risk in recessions and expansions. *The Quarterly Journal of Finance*, 1(3):607–664, 2011.

[47] Martin Lettau and Sydney C. Ludvigson. Measuring and modeling variation in the risk-return tradeoff. In *Handbook of Financial Econometrics*. Ait-Sahalia and Hansen, L. P. Amsterdam: North-Holland, 2010.

[48] Sydney C. Ludvigson and Serena Ng. Macro factors in bond risk premia. *Review of Financial Studies*, 22(12):5027–5067, 2009.

[49] Bartosz Maćkowiak and Mirko Wiederholt. Optimal sticky prices under rational inattention. *The American Economic Review*, 99(3):769–803, 2009.

[50] Bartosz Maćkowiak and Mirko Wiederholt. Business cycle dynamics under rational inattention. *The Review of Economic Studies*, page rdv027, 2015.

[51] Katya Malinova, Andreas Park, and Ryan Riordan. Do retail traders suffer from high frequency traders? *Available at SSRN*, 2016.

[52] Luboš Pástor and Robert F Stambaugh. Investing in equity mutual funds. *Journal of Financial Economics*, 63(3):351–380, 2002.

[53] Ryan Riordan and Adreas Storkenmaier. Latency, liquidity, and price discovery. *Journal of Financial Markets*, 15(4):416–437, 2012.

[54] Andriy Shkilko and Konstantin Sokolov. Every cloud has a silver lining: Fast trading, microwave connectivity and trading costs. *Available at SSRN*, 2016.

[55] Christopher A. Sims. Implications of rational inattention. *Journal of monetary Economics*, 50(3):665–690, 2003.

[56] Christopher A. Sims. Rational inattention: Beyond the linear-quadratic case. *The American economic review*, 96(2):158–163, 2006.

[57] Hans R. Stoll. High speed equities trading: 1993-2012, asia-pacific. *Journal of Financial Studies*, 43(6):767–797, 2014.

[58] Stijn Van Nieuwerburgh and Laura Veldkamp. Information acquisition and under-diversification. *The Review of Economic Studies*, 77(2):779–805, 2010.

APPENDIX

A Fund Classification

1. The words utilized in the "word search" of fund behaviors in mutual fund prospectuses is summarized as follows (variations of each word/clause were considered):

(a) activeTrading - active trading, actively trade, actively traded, actively and frequently traded, actively managed

(b) frequentTrading - short-term trading, short-term trade, shortterm trade, frequent trading, frequently traded, frequently and actively traded, frequent and active trading

(c) shortsell - short-selling, short-sell, selling securities short, sell securities short, selling short, short portfolio, sell short, short-sale, sold short, shorting, sells, short position, 110/10, 115/15, 130/30, 150/50, long-short, long/short strategy, short positions, short position, sells a security short

(d) trendFollowing - behavioral, sentiment, momentum, behavior factors, momentum/sentiment, technical analysis, chart patterns, uptrend, up-trending, stock price up-trending, price downtrend, down-trending, relative price down-trending, price changes in trend direction, analyzing long-term relative price trends, analyzing price trends, analysis of price trends, investor psychology, long term trends, relative price movement

B Useful Derivations

1. The sensitivity of the marginal benefit of learning about risk factor i with respect to its expected supply is positive and is given by:

$$\frac{\partial \lambda_i}{\partial \bar{x}_i^2} = \frac{\partial}{\partial \bar{x}_i^2}\left[\bar{\sigma}_i + \bar{\sigma_i}^2\rho^2\sigma_x + \bar{\sigma_i}^2\rho^2\sigma_x\bar{x}_i^2 + \bar{\sigma_i}^2\bar{K}_i\right] = \bar{\sigma_i}^2\rho^2 > 0 \tag{33}$$

2. The sensitivity of the marginal benefit of learning about risk factor i with respect to its volatility is positive and is given by:

$$\frac{\partial \lambda_i}{\partial \sigma_i} = \left\{1 + 2\left[\rho^2\left(\sigma_x + \bar{x}_i^2\right) + \bar{K}_i\right]\sigma_i\right\}\left(\frac{\bar{\sigma}_i}{\sigma_i}\right)^2 > 0 \tag{34}$$

This positive effect is decreasing in the share of skilled funds who use quantitative strategies θ for $i = 1, ..., n-1$, for shocks with a positive information gap:

$$\frac{\partial \lambda_i}{\partial \sigma_i \partial \theta} = -2\chi\left(\sigma_{\eta i}^{-1} - \bar{K}_{id}\right)\bar{\sigma}_i\left\{3\bar{\sigma}_i\left[\rho^2\left(\sigma_x + \bar{x}_i^2\right) + \bar{K}_i\right]\left(\frac{2\bar{K}_i}{\rho^2\sigma_x} + 1\right) + \frac{2\bar{K}_i}{\rho^2\sigma_x}\right\}\left(\frac{\bar{\sigma}_i}{\sigma_i}\right)^2 \tag{35}$$

It is instead always increasing for $i = n$, when $\sigma_{\eta i}^{-1} = 0$:

$$\frac{\partial \lambda_i}{\partial \sigma_i \partial \theta} = 2\chi\bar{K}_{id}\bar{\sigma}_i\left\{3\bar{\sigma}_i\left[\rho^2\left(\sigma_x + \bar{x}_i^2\right) + \bar{K}_i\right]\left(\frac{2\bar{K}_i}{\rho^2\sigma_x} + 1\right) + \frac{2\bar{K}_i}{\rho^2\sigma_x}\right\}\left(\frac{\bar{\sigma}_i}{\sigma_i}\right)^2 \tag{36}$$

3. The sensitivity of the marginal benefit of learning about risk factor i with respect to the average precision of private signals about i ($\bar{K}_i$) is negative and is given by

$$\frac{\partial \lambda_i}{\partial \bar{K}_i} = -2\bar{\sigma}_i\left\{\frac{\bar{K}_i}{\rho^2\sigma_x} + \bar{\sigma}_i\left[\rho^2\left(\sigma_x + \bar{x}_i^2\right) + \bar{K}_i\right]\left(\frac{2\bar{K}_i}{\rho^2\sigma_x} + 1\right)\right\} < 0 \tag{37}$$

4. The sensitivity of the marginal benefit of learning about risk factor i with respect to the share of skilled investors using quantitative strategies is given by:

$$\frac{\partial \lambda_i}{\partial \theta} = -2\chi \left(\sigma_{\eta i}^{-1} - K_{id}\right) \bar{\sigma}_i^2 \left\{ \frac{\bar{K}_i}{\rho^2 \sigma_x} \left[1 + 2\bar{\sigma}_i \left[\rho^2 \left(\sigma_x + \bar{x}_i^2\right) + \bar{K}_i\right]\right] + \bar{\sigma}_i \left[\rho^2 \left(\sigma_x + \bar{x}_i^2\right) + \bar{K}_i\right] \right\} \tag{38}$$

Exploring the sign of the derivative in equation 38 we can see that:

$$\begin{cases} \frac{\partial \lambda_i}{\partial \theta} < 0 & i = 1, ..., n-1 \ and \ G_i > 0 \\ \frac{\partial \lambda_i}{\partial \theta} > 0 & i = 1, ..., n-1 \ and \ G_i < 0 \\ \frac{\partial \lambda_i}{\partial \theta} > 0 & i = n \end{cases} \tag{39}$$

For $i = n$, $\sigma_{\eta n}^{-1} = 0$, for all other risk factors instead we have that a rise in $\sigma_{\eta i}^{-1}$ reduces the positive effect ($G_i < 0$) or increases the negative ($G_i < 0$) effect of θ on λ_i:

$$\frac{\partial \lambda_i}{\partial \theta \partial \sigma_{\eta i}^{-1}} = -2\chi \bar{\sigma}_i^2 \{\chi\theta \left(\sigma_{\eta i}^{-1} - K_{id}\right) \left[\frac{1}{\rho^2 \sigma_x} + \bar{\sigma}_i + \frac{2\bar{\sigma}_i}{\rho^2 \sigma_x} \left[2\bar{K}_i + \rho^2 \left(\sigma_x + \bar{x}_i^2\right)\right]\right] + \frac{\bar{K}_i}{\rho^2 \sigma_x} + \bar{\sigma}_i \bar{K}_i + \bar{\sigma}_i \rho^2 \left(\sigma_x + \bar{x}_i^2\right)\} < 0 \tag{40}$$

A similar reasoning applies for increases in the information gap G_i:

$$\frac{\partial \lambda_i}{\partial \theta \partial G_i} = -2\chi \bar{\sigma}_i^2 \{\chi\theta G_i \left[\frac{1}{\rho^2 \sigma_x} + \bar{\sigma}_i + \frac{2\bar{\sigma}_i}{\rho^2 \sigma_x} \left[2\bar{K}_i + \rho^2 \left(\sigma_x + \bar{x}_i^2\right)\right]\right] + \frac{\bar{K}_i}{\rho^2 \sigma_x} + \bar{\sigma}_i \bar{K}_i + \bar{\sigma}_i \rho^2 \left(\sigma_x + \bar{x}_i^2\right)\} + \frac{\bar{K}_i}{\rho^2 \sigma_x} + \bar{\sigma}_i \bar{K}_i + \bar{\sigma}_i \rho^2 (\sigma \tag{41}$$

Finally also increases in the expected supply of risk factor i reduces the positive effect ($G_i < 0$) or increases the negative ($G_i < 0$) effect of θ on λ_i:

$$\frac{\partial \lambda_i}{\partial \theta \partial \bar{x}_i^2} = -2\chi \left(\sigma_{\eta i}^{-1} - K_{id}\right) \bar{\sigma}_i^3 \left(\frac{2\bar{K}_i}{\sigma_x} + \rho^2\right) \tag{42}$$

Exploring the sign of the derivative in equation 42 we can see that:

$$\begin{cases} \frac{\partial \lambda_i}{\partial \theta \partial \bar{x}_i^2} < 0 & i = 1, ..., n-1 \ and \ G_i > 0 \\ \frac{\partial \lambda_i}{\partial \theta \partial \bar{x}_i^2} > 0 & i = 1, ..., n-1 \ and \ G_i < 0 \\ \frac{\partial \lambda_i}{\partial \theta \partial \bar{x}_i^2} > 0 & i = n \end{cases}$$

5. The elasticities of $\lambda_{i \neq n}$ and λ_n with respect to θ are given by:

$$e_{\lambda_{i \neq n}, \theta} = -\frac{\partial \lambda_{i \neq n}}{\partial \theta} \frac{\theta}{\lambda_{i \neq n}} > 0; \ e_{\lambda_n, \theta} = \frac{\partial \lambda_n}{\partial \theta} \frac{\theta}{\lambda_n} > 0 \tag{43}$$

C Model Solution

C.1 *Market-clearing*

The model is solved in terms of synthetic assets only affected by one risk-factor each, after the model has been solved, we can map the results back from synthetic assets to real ones.

Solving for the market clearing condition we obtain that:

$$A = \Gamma^{-1}\mu - \rho\Sigma\bar{x};\ B = I - \bar{\Sigma}\Sigma^{-1};\ C = -\rho\bar{\Sigma}\left(I - \frac{1}{\rho^2\sigma_x}\bar{\Sigma}_\eta^{-1}\right)$$

$$\tilde{pr} = \Gamma^{-1}\mu + \Sigma\left[\left(\bar{\Sigma}^{-1} - \Sigma^{-1}\right)z - \rho(\bar{x} - x) - \frac{1}{\rho^2\sigma_x}\bar{\Sigma}_\eta^{-1}x\right]$$

Where:

$$\bar{\Sigma}^{-1} = \Sigma^{-1} + \Sigma_p^{-1} + \bar{\Sigma}_\eta^{-1};\ \Sigma_P^{-1} = \left(\sigma_x B^{-1}CC'B^{-1'}\right)^{-1} = \frac{1}{\rho^2\sigma_x}\bar{\Sigma}_\eta^{-1'}\bar{\Sigma}_\eta^{-1}$$

See Admati (1985) and Kacperczyk, Van Nieuwerburgh, and Veldkamp (2016) for more details on the derivation.

C.2 *Discretionary learning choice*

The discretionary investors' optimal learning problem is obtained by solving the following system of equations:

$$\begin{cases} U_{1d} = E_1\left[\rho E_d\left[W_d\right] - \frac{\rho^2}{2}V_d\left[W_d\right]\right] \\ W_d = rW_0 + \tilde{q}_d^{*'}\left(\tilde{f} - \tilde{pr}\right) \\ \tilde{q_d}^* = \frac{1}{\rho}\hat{\Sigma_d}^{-1}\left(E_d[\tilde{f}] - \tilde{pr}\right) \end{cases}$$

The solution is obtained by following the steps below:

1. Substitute $\tilde{q_d}^*$ into W_d
2. Compute $E_d\left[W_d|\hat{E}_d\left[\tilde{f}\right], \hat{\Sigma}_d\right]$ and $V_d\left[W_d|\hat{E}_d\left[\tilde{f}\right], \hat{\Sigma}_d\right]$
3. Substitute the conditional moments derived at point 2 into the equation for U_{1d}

See Kacperczyk, Van Nieuwerburgh, and Veldkamp (2016) for more details on the derivation.

D Learning choice of discretionary skilled investors

D.1 *Proofs of Propositions 1 – Part 1 – and Proposition 3*

As can be seen from equation 39, an increase in θ increases the marginal benefit of learning about the aggregate shock (λ_n) and about idiosyncratic shocks with a negative information gap ($\lambda_{i\neq n}$ *and* $G_i < 0$) and decreases the marginal benefit of learning about idiosyncratic shocks with a positive information gap ($\lambda_{i\neq n}$ *and* $G_i > 0$). In order to verify the effect of increases in θ on the attention allocation choice of discretionary investors we need to also consider how attention was allocated before the increase. For this reason we need to consider three cases:

- CASE 1: Before the increase in θ no attention was allocated to λ_i. For $i = n$ and for $i \neq n$ *and* $G_i < 0$, λ_i increases but λ_i is continuous in θ so the marginal increase in θ cannot change the ranking of $\lambda_i s$ so we don't optimally observe any change in the attention allocation. For $i \neq n$ *and* $G_i > 0$, λ_i decreases, reducing even further the incentive to learn about it, so the optimal attention allocation is not affected.

- CASE 2: Before the increase in θ all attention was already allocate to λ_i. For $i = n$ and for $i \neq n$ *and* $G_i < 0$, λ_i increases even further but being all attention already allocated to it, optimal attention allocation doesn't change. For $i \neq n$ *and* $G_i > 0$, λ_i decreases but being λ_i continuous in θ, a marginal increase in θ cannot change the discrete ranking of $\lambda_i s$ so attention allocation doesn't change.

- CASE: Before the increase in θ, λ_i was within the basket of assets among which attention was allocated. In this case the equilibrium solution is obtained by waterfilling (see Cover and Thomas 1991; Kacperczyk, Van Nieuwerburgh, and Veldkamp 2014).

  - CASE 3a: For $i = n$, an increase in θ increases λ_n, this increases the attention allocated to the risk factor n (σ_{dn}^{-1}). The increase in attention allocation, though, decreases the marginal benefit of learning about n (λ_n) due to substitution effects (as shown in Appendix B). Additionally an increase in θ also decreases the marginal benefit of learning about idiosyncratic risk factors in the basket with a positive information gap ($\lambda_{i \neq n}, G_i > 0$), which decrease the attention allocated to them (σ_{di}^{-1}). This decrease in attention in turn increases the marginal benefit of learning about these factors. Finally an increase in θ will also increase the marginal benefit of learning about idiosyncratic shocks in the basket with a negative information gap ($\lambda_{i \neq n}, G_i < 0$); but this increase will be relatively lower than the increase in the marginal benefit of learning about the aggregate shock. This is because the marginal benefit of learning about a shock is increasing in its expected supply and the aggregate shock is in greatest supply. Through the waterfilling methodology, we reallocate attention among the risk factors in the basket until either $\lambda_n = argmax_i \lambda_i$ and all attention is allocated to the aggregate risk factor; or the marginal benefit of learning is again equalized among all risk factors in the basket. In both cases, though, the attention allocated to the aggregate risk factor strictly increases.
  - CASE 3b: For $i \neq n$ and $G_i > 0$ (resp.$G_i < 0$), an increase in θ decreases (resp. increases) λ_i, this decreases (resp. increases) the attention allocated to risk factor i (σ_{di}^{-1}). The decrease (resp. increase) in attention allocation, in turn, increases (resp. decreases) the incentive to learn about risk factor i, due to substitution effects. Through waterfilling, we continue reallocating attention among the risk factors in the basket until either no (resp. all) attention is allocated to risk factor i or the marginal benefit of learning about the risk factors in the basket is again equalized. An increase in θ, though, has a negative (resp. positive) effect on ALL idiosyncratic risk factors in the basket. We can see, though, that $\frac{\partial \lambda_i}{\partial \theta}$ is strictly decreasing in G_i. This means that an increase in θ will have a larger

effect in decreasing (resp. increasing) the marginal benefit of learning about those risk factors with a greater (resp. smaller) information gap. Therefore there exists G_i^* such that $\frac{\partial\lambda_i}{\partial\theta} < \frac{\partial\lambda_i^{'}}{\partial\theta}\forall i^{'}$ (resp. $\frac{\partial\lambda_i}{\partial\theta} > \frac{\partial\lambda_i^{'}}{\partial\theta}\forall i^{'}$) if $G_i < G_i^*$. Hence in this case we will observe a shift of attention among the risk factors in the basket towards those risk factors with a relatively smaller information gap.

To summarize, in cases one and two we don't observe any change in the optimal attention allocation. In case three, if attention was already partially allocated to the aggregate risk factor we observe an increase in the attention allocated to it. On the other hand if attention was partially allocated to a basket of idiosyncratic risk factors we observe a shift of attention towards those risk factors with a relatively smaller information gap. Hence an increase in θ weakly increases the attention towards the aggregate shock and weakly decreases the attention towards idiosyncratic shocks with a relatively greater information gap.

D.2 *Proof of Proposition 1 - Part 2*

There exist values of $\theta \in [0, 1]$ high enough for which in expansions, whenσ_n and ρ are low, discretionary skilled investors still choose to specialize in learning about the aggregate risk factor ($\lambda_n^E > \lambda_{i\neq n}^E$). By analyzing the common terms in the equations for λ_n^E and $\lambda_{i\neq n}^E$ below:

$$\lambda_{i\neq n}^E = \frac{1}{\sigma_{i\neq n}^{-1} + \chi\theta\sigma_{\eta i\neq n}^{-1} + \chi(1-\theta)K + \frac{(\chi\theta\sigma_{\eta i\neq n}^{-1}+\chi(1-\theta)K)^2}{\rho^2\sigma_x}} + \frac{\rho^2\sigma_x + +\chi(1-\theta)K + \rho^2\bar{x}_{i\neq n}^2 + \chi\theta\sigma_{\eta i\neq n}^{-1}}{\left(\sigma_{i\neq n}^{-1} + \chi\theta\sigma_{\eta i\neq n}^{-1} + \chi(1-\theta)K + \frac{(\chi\theta\sigma_{\eta i\neq n}^{-1}+\chi(1-\theta)K)^2}{\rho^2\sigma_x}\right)^2}$$

$$\lambda_n^E = \frac{1}{\sigma_n^{-1} + \chi(1-\theta)K + \frac{(\chi(1-\theta)K)^2}{\rho^2\sigma_x}} + \frac{\rho^2\sigma_x + \chi(1-\theta)K + \rho^2\bar{x}_i{}^2}{\left(\sigma_i^{-1} + \chi(1-\theta)K + \frac{(\chi(1-\theta)K)^2}{\rho^2\sigma_x}\right)^2}$$

We can see that a restrictive sufficient condition for λ_n^E to be greater than $\lambda_{i\neq n}^E$ is that there exist values of $\theta < 1$ such that: $\sigma_n^{-1} < \sigma_{i\neq n}^{-1} + \chi\theta\sigma_{\eta i\neq n}^{-1}$ even when $\sigma_n <\sigma_{i\neq n}$. Or equivalently: $\sigma_{\eta i\neq n}^{-1} > \frac{\sigma_{i\neq n}-\sigma_n}{\chi\sigma_{i\neq n}\sigma_n}$. Hence for sufficiently high precision of the quantitative funds signal there exist values of $\theta < 1$ such that discretionary skilled funds would always choose to specialize in learning about the aggregate shock. To put this into context let's assume that only 20% of all investors in the market are skilled $\chi = 0.2$ and that the volatility of the idiosyncratic risk of choice in expansions is double that of the aggregate shock (eg $\sigma_i^* = 0.5$, $\sigma_n = 0.25$). Then the condition above becomes: $\sigma_{\eta i\neq n} < \frac{\theta}{5}\sigma_{i\neq n}$ from which we can see that for a low $\theta = 2\%$ in order for the condition to be verified we would need the precision of the quants signal to be incredibly high $\sigma_{\eta i\neq n} < 0.4\%\sigma_i$ (which for a $\sigma_i = 0.5$ would mean a precision of the quants' signal of $\sigma_{\eta i\neq n}^{-1} = 500$). If though considering that the skilled investors were evenly divided among quantitative and discretionary $\theta = 50\%$ in order for the condition to be verified we would need the precision of the quant signal to be only $\sigma_{\eta i\neq n} < 10\%\sigma_i$(which for a $\sigma_i = 0.5$ would mean a precision of the quants' signal of $\sigma_{\eta i\neq n}^{-1} = 20$).

Hence, as the share of quantitative skilled investors in the market rises, the incentive for specialization rises, for progressively lower levels of precision of the quant signal.

D.3 *Proof of Proposition 4*

Let us define the risk factor portfolio of the average investor as:

$$\int \tilde{q}_j^* \, dj = \frac{1}{\rho} \int \hat{\Sigma}_j^{-1} \left(E_j[\tilde{f}] - \tilde{p}r \right) \partial j = \frac{1}{\rho} \int \left[\hat{\Sigma}_j^{-1} \left(\Gamma' \mu + E_j[z] \right) \right] dj - \bar{\Sigma}_j^{-1} \tilde{p}r$$

$$= \frac{1}{\rho} \left\{ \int \left[\hat{\Sigma}_j^{-1} E_j[z] \right] dj + \bar{\Sigma}^{-1} \left(\Gamma' \mu - \tilde{p}r \right) \right\} = \frac{1}{\rho} \left\{ \int \left[\hat{\Sigma}_j^{-1} \hat{\Sigma}_j \left(\Sigma_{\eta j}^{-1} \eta_j + \Sigma_p^{-1} \eta_p \right) \right] dj + \bar{\Sigma}_j^{-1} \left(\Gamma' \mu - \tilde{p}r \right) \right\}$$

$$\int \tilde{q}_j^* \, \partial j = \frac{1}{\rho} \left[\Sigma_{\eta j}^{-1} z + \Sigma_p^{-1} \eta_p + \bar{\Sigma}_j^{-1} \left(\Gamma' \mu - \tilde{p}r \right) \right]$$

Similarly the risk factor portfolio for the average quantitative and discretionary investors are given respectively by:

$$\int \tilde{q}_q^* \, \partial q = \frac{1}{\rho} \left[\Sigma_{\eta q}^{-1} z + \Sigma_p^{-1} \eta_p + \hat{\Sigma}_q^{-1} \left(\Gamma' \mu - \tilde{p}r \right) \right] = \bar{\tilde{q}}_q$$

$$\int \tilde{q}_d^* \, \partial d = \frac{1}{\rho} \left[\bar{\Sigma}_{\eta d}^{-1} z + \Sigma_p^{-1} \eta_p + \bar{\Sigma}_d^{-1} \left(\Gamma' \mu - \tilde{p}r \right) \right] = \bar{\tilde{q}}_d$$

where $\bar{\Sigma}_d^{-1} = \Sigma^{-1} + \Sigma_p^{-1} + \int \Sigma_{\eta d}^{-1} \, \partial d$ and $\bar{\Sigma}_q^{-1} = \Sigma^{-1} + \Sigma_p^{-1} + \int \Sigma_{\eta q}^{-1} \, \partial q = \Sigma^{-1} + \Sigma_p^{-1} + \Sigma_{\eta q}^{-1} = \hat{\Sigma}_q^{-1}$. The last equation comes from the fact that each quantitative investors optimally drives the volatility of their private signals to the lower-bound for each idiosyncratic shock and they all have zero precision for the aggregate shock signal, hence they all have the same posterior precision about all shocks.

Then, the difference between the risky portfolio of a quantitative or discretionary investor and that of the average quantitative or discretionary investor is respectively given by:

$$(\tilde{q}_q^* - \bar{\tilde{q}}_q) = \frac{1}{\rho} \left[\hat{\Sigma}_q^{-1} \left(E_q[\tilde{f}] - \tilde{p}r \right) - \Sigma_{\eta q}^{-1} z + \Sigma_p^{-1} \eta_p + \hat{\Sigma}_q^{-1} \left(\Gamma' \mu - \tilde{p}r \right) \right] = \frac{1}{\rho} \Sigma_{\eta q}^{-1} \epsilon_q$$

$$(\tilde{q}_d^* - \bar{\tilde{q}}_d) = \frac{1}{\rho} \left[\hat{\Sigma}_d^{-1} \left(E_d[\tilde{f}] - \tilde{p}r \right) - \bar{\Sigma}_{\eta d}^{-1} z + \Sigma_p^{-1} \eta_p + \bar{\Sigma}_d^{-1} \left(\Gamma' \mu - \tilde{p}r \right) \right] = \frac{1}{\rho} \left[\left(\Sigma_{\eta d}^{-1} - \bar{\Sigma}_{\eta d}^{-1} \right) \left(\tilde{f} - \tilde{p}r \right) + \Sigma_{\eta d}^{-1} \epsilon_d \right]$$

where the difference in the case of the discretionary investor includes an additional term, deriving from the potentially different precision of the signal of each discretionary investor from that of the average discretionary investor. This derives from the fact that they all have to optimally choose where to allocate their capacity before deciding on their optimal allocation, hence they might choose to learn about different shocks.

Dispersion of opinion is then defined as the expected square difference between the risky portfolio of an investor and that of the average investor in his group.

Dispersion of opinion for a quantitative investor is then simply given by:

$$E\left[\left(\tilde{q}_q^* - \bar{\tilde{q}}_q\right)\left(\tilde{q}_q^* - \bar{\tilde{q}}_q\right)'\right] = \frac{1}{\rho^2}\sum_{i=1}^{n-1} \underline{\sigma}_{iq}^{-1}$$

The dispersion of opinion of discretionary investors has to also take into account the difference in signal precision among discretionary investors. It is then given by:

$$E\left[\left(\tilde{q}_d^* - \bar{\tilde{q}}_d\right)\left(\tilde{q}_d^* - \bar{\tilde{q}}_d\right)'\right] = \frac{1}{\rho^2} E\left[\sum_{i=1}^{n}\left(\sigma_{id}^{-1} - K_{id}\right)\left(\tilde{f} - \tilde{p}r\right) + \sigma_{id}^{-1}\epsilon_{id}\right]$$

$$= \frac{1}{\rho^2}\left[\sum_{i=1}^{n}\left(\sigma_{id}^{-1} - K_{id}\right)\left(V_{ii} + \rho^2\bar{\sigma}_i^2\bar{x}_i{}^2\right)\right] + \frac{1}{\rho^2}K$$

Dispersion of opinion is greater among discretionary investors than among quantitative investors if for any quantitative investor (q) and discretionary investor (d), the discretionary investor has greater dispersion in his risky portfolio with respect to the average discretionary investor than quantitative investors do. Hence if for each q and d the following condition is verified:

$$\left[\sum_{i=1}^{n}\left(\sigma_{id}^{-1} - K_{id}\right)\left(V_{ii} + \rho^2\bar{\sigma}_i^2\bar{x}_i{}^2\right)\right] > \sum_{i=1}^{n-1} \underline{\sigma}_{iq}^{-1} - K$$

E Type-conditional over-performance

E.1 *Derivation of expected excess returns*

The expected excess return is obtained by following steps below:

1. Solve for $(\tilde{f} - \tilde{p}r)$
2. Solve for $(\tilde{q}_j - \bar{\tilde{q}})$
3. Substitute them into: $(q_j - \bar{q})'(f - pr) = (q_j - \bar{q})'\mathit{\Gamma}^{-1}(\mathit{\Gamma} f - \mathit{\Gamma} pr) = (\tilde{q}_j - \bar{\tilde{q}})'(\tilde{f} - \tilde{p}r)$
4. Take the expectation: $E[(\tilde{q}_j - \bar{\tilde{q}})'(\tilde{f} - \tilde{p}r)] = \rho Tr(\bar{x}'\bar{\Sigma}\Delta\bar{\Sigma}\bar{x}) + \frac{1}{\rho}Tr(\Delta V) = \frac{1}{\rho}\sum_{i=1}^{n}\left[\lambda_i\left(K_{ij} - \bar{K}_i\right)\right]$

See Kacperczyk, Van Nieuwerburgh, and Veldkamp (2016) for more details on the derivation.

E.2 *Proof of Proposition 5*

The expected excess return of skilled quantitative investors is given by:

$$E[R_Q - R_M] = \frac{1}{\rho}\sum_{i=1}^{n-1}\left[\lambda_i\left(\underline{\sigma}_{iq}^{-1}(1-\chi\theta) - K_{id}\chi(1-\theta)\right)\right] - \frac{1}{\rho}\lambda_n K_{nd}\chi(1-\theta)$$

Exploring the sign:

- When $K_{nd} = 0$ (Discretionaries are not learning about the aggregate shock), we have that:

$$E[R_Q - R_M] = \frac{1}{\rho}\sum_{i=1}^{n-1}\left[\lambda_i\left(\underline{\sigma}_{iq}^{-1}(1-\chi\theta) - K_{id}\chi(1-\theta)\right)\right] \tag{44}$$

which is always greater than zero as most stocks have a positive information gap.

- When $K_{nd} = K$ (Discretionaries are learning about the aggregate shock) we have that:

$$E[R_Q - R_M] = \frac{1}{\rho}\sum_{i=1}^{n-1}\left[\lambda_i\left(\underline{\sigma}_{iq}^{-1}(1-\chi\theta)\right)\right] - \frac{1}{\rho}\lambda_n K\chi(1-\theta) \tag{45}$$

This equation is greater than zero when:

$$\frac{1}{\rho}\sum_{i=1}^{n-1}\left[\lambda_i\left(\underline{\sigma}_{iq}^{-1}(1-\chi\theta)\right)\right] > \frac{1}{\rho}\lambda_n K\chi(1-\theta) \textit{ implies } (1-\chi\theta)\sum_{i=1}^{n-1}\left[\lambda_i\underline{\sigma}_{iq}^{-1}\right] > (\chi-\chi\theta)\lambda_n K$$

$$\sum_{i=1}^{n-1}\left[\lambda_i\underline{\sigma}_{iq}^{-1}\right] - \chi\theta\sum_{i=1}^{n-1}\left[\lambda_i\underline{\sigma}_{iq}^{-1}\right] > \chi\lambda_n K - \chi\theta\lambda_n K \textit{ implies } \sum_{i=1}^{n-1}\left[\lambda_i\underline{\sigma}_{iq}^{-1}\right] - \chi\lambda_n K > \chi\theta\sum_{i=1}^{n-1}\left[\lambda_i\underline{\sigma}_{iq}^{-1}\right] - \chi\theta\lambda_n K$$

$$\frac{\sum_{i=1}^{n-1}\left[\lambda_i\underline{\sigma}_{iq}^{-1}\right] - \chi\lambda_n K}{\chi\sum_{i=1}^{n-1}\left[\lambda_i\underline{\sigma}_{iq}^{-1}\right] - \chi\lambda_n K} > \theta \textit{ implies } \frac{\sum_{i=1}^{n-1}\left[\lambda_i\underline{\sigma}_{iq}^{-1}\right] - \chi\lambda_n K}{\chi\sum_{i=1}^{n-1}\left[\lambda_i\underline{\sigma}_{iq}^{-1}\right] - \chi\lambda_n K} > \theta \tag{46}$$

The above is always verified as the left hand side of equation 46 is always greater than 1 and $\theta \in [0, 1]$.

Effect of increases in σ_i:

- When $K_{nd} = 0$ and $K_{jd} = K$ for $j \neq n$ (Without loss of generality discretionaries are learning about idiosyncratic shock j - which can be considered as a single risk factor or a basket of factors):

$$E[R_Q - R_M] = \frac{1}{\rho}\sum_{i\neq j}^{n-1}\left[\lambda_i\left(\underline{\sigma}_{iq}^{-1}(1-\chi\theta)\right)\right] + \frac{1}{\rho}\lambda_j\left(\underline{\sigma}_{jq}^{-1}(1-\chi\theta) - K\chi(1-\theta)\right)$$

An increase in σ_n has no effect. An increase in $\sigma_{i \neq j}$ only affects λ_i ($\frac{\partial \lambda_i}{\partial \sigma_i} > 0$), hence $\frac{\partial E[R_Q - R_M]}{\partial \sigma_i} > 0$. Finally an increase in σ_j only affects λ_j ($\frac{\partial \lambda_j}{\partial \sigma_j} > 0$), since $\underline{\sigma}_{jq}^{-1}(1-\chi\theta) - K\chi(1-\theta)$ is always positive ($K_{jq} > K$), this has a positive impact on excess return: $\frac{\partial E[R_Q - R_M]}{\partial \sigma_j} > 0$.

- When $K_{nd} = K$ (Discretionaries are learning about the aggregate shock):

$$E[R_Q - R_M] = \frac{1}{\rho}\sum_{i=1}^{n-1}\left[\lambda_i\left(\underline{\sigma}_{iq}^{-1}(1-\chi\theta)\right)\right] - \frac{1}{\rho}\lambda_n K\chi(1-\theta)$$

An increase in σ_n only affects λ_n ($\frac{\partial \lambda_n}{\partial \sigma_n} > 0$), hence $\frac{\partial E[R_Q - R_M]}{\partial \sigma_n} < 0$. Whereas an increase in σ_i only affects λ_i ($\frac{\partial \lambda_i}{\partial \sigma_i} > 0$), hence $\frac{\partial E[R_Q - R_M]}{\partial \sigma_i} > 0$.

Effect of increases in θ:

$$E[R_Q - R_M] = \frac{1}{\rho}\sum_{i \neq j}^{n-1}\left[\lambda_i\left(\underline{\sigma}_{iq}^{-1}(1-\chi\theta)\right)\right] + \frac{1}{\rho}\sum_{j=1}^{J}\lambda_j\left(\underline{\sigma}_{jq}^{-1}(1-\chi\theta) - K\chi(1-\theta)\right) - \frac{1}{\rho}\lambda_n K\chi 1 + \frac{1}{\rho}\lambda_n K\chi\theta$$

$$\frac{\partial E[R_Q - R_M]}{\partial \theta} = \frac{1}{\rho}\left\{\sum_{i \neq j}^{n-1}\frac{\partial \lambda_i}{\partial \theta}\underline{\sigma}_{iq}^{-1}(1-\chi\theta) + \lambda_i\frac{\partial}{\partial \theta}\left[\underline{\sigma}_{iq}^{-1}(1-\chi\theta)\right]\right\} +$$

$$+\frac{1}{\rho}\left\{\sum_{j=1}^{J}\frac{\partial \lambda_j}{\partial \theta}\left[\underline{\sigma}_{iq}^{-1} - \chi K_{id} - \chi\theta(\underline{\sigma}_{iq}^{-1} - K_{id})\right] + \lambda_j\frac{\partial}{\partial \theta}\left[\underline{\sigma}_{iq}^{-1} - \chi K_{id} - \chi\theta(\underline{\sigma}_{iq}^{-1} - K_{id})\right]\right\} - \frac{1}{\rho}\frac{\partial \lambda_n}{\partial \theta}\chi K + \frac{\partial \lambda_n}{\partial \theta}\chi\theta K + \lambda_n\frac{\partial}{\partial \theta}\left[K\chi\theta\right]$$

Recalling from appendix B the elasticities definitions and solving them for $\frac{\partial \lambda_n}{\partial \theta}$ and $\frac{\partial \lambda_i}{\partial \theta}$ after substitution and some manipulation we obtain:

$$\frac{\partial E[R_Q - R_M]}{\partial \theta} = \frac{1}{\rho}\left\{\sum_{i \neq j}^{n-1}\left[\underline{\sigma}_{iq}^{-1}\lambda_i\left(e_{\lambda_i,\theta}\left(\chi - \frac{1}{\theta}\right) - \chi\right)\right]\right\} + \tag{47}$$

$$+\frac{1}{\rho}\left\{\sum_{j=1}^{J}\left[\lambda_j\left(e_{\lambda_j,\theta}\left(\underline{\sigma}_{jq}^{-1}\left(\chi - \frac{1}{\theta}\right) - K_{j,d}\left(\chi^2 - \frac{1}{\theta}\right)\right) - \chi\left(\underline{\sigma}_{jq}^{-1} - K_{jd}\right)\right)\right]\right\} + \tag{48}$$

$$+\frac{1}{\rho}\left\{\chi K\left[\lambda_n(1 + e_{\lambda_n,\theta}\left(1 - \frac{1}{\theta}\right)\right]\right\} \tag{49}$$

Analyzing the sign of the equation above we can see that the first part of the equation (equation number 47) is clearly negative as all elements are positive whereas $\left(\chi - \frac{1}{\theta}\right)$ is negative. The equation becomes more negative the greater is the elasticity of λ_i to θ ($e_{\lambda_i,\theta}$), which acts as a multiplier. In other words, as the negative impact of θ on λ_i increases, further increases in θ have a greater negative impact on the excess return of the quants as the marginal benefit of learning about risk factor i

decreases at a higher rate (overcrowding happens faster). Finally also the precision of the quants signal acts as a multiplier; hence the higher the precision $\underline{\sigma}_{iq}^{-1}$ the larger is the negative impact of increases in θ on the excess return of the quantitative skilled investors.

Now looking at the second part of the above equation (equation number 48) we can see that it is also negative if the precision of the quants signal is sufficiently larger than that of the discretionaries ($\underline{\sigma}_{iq}^{-1} \gg K_{jd}$ – greater information gap), with a similar reasoning, the impact of increases of θ on the excess return of the quantitative skilled investors is more negative the greater is the elasticity of λ_j to θ ($e_{\lambda_j,\theta}$) and the greater is $\underline{\sigma}_{jq}^{-1}$.

Finally, looking at the third part of the above equation (equation number 49) we can see that again it is negative if the elasticity of λ_n to θ ($e_{\lambda_n,\theta}$) is sufficiently large.

As a summary, increases in the share of skilled investors who utilize quantitative strategies has a greater negative impact on the excess return of quantitative skilled investors the greater is the precision of the quantitative investors' private signal and the greater is the elasticity of the marginal benefit of learning about the different risk factors to θ. This is due to overcrowding; indeed the greater is the informational content in the market (due to the share of informed investors and how much each of them knows) the lower is the benefit of learning due to substitution effects.

E.3 *Proof of Proposition 6*

The expected excess return of the skilled discretionary investors is given by:

$$E[R_D - R_M] = \frac{1}{\rho}\sum_{i=1}^{n-1}\left\{\lambda_i\left[\sigma_{id}^{-1} - K_{id}(\chi(1-\theta)) - \underline{\sigma}_{iq}^{-1}\chi\theta\right]\right\} + \frac{1}{\rho}\lambda_n\sigma_{nd}^{-1} - K_{nd}(\chi(1-\theta)) \tag{50}$$

Exploring the sign

- When choosing to learn about the aggregate shock $\sigma_{nd}^{-1} = K$:

$$E[R_D - R_M] = \frac{1}{\rho}\left[\lambda_n K - (\chi(1-\theta))\lambda_n K_{nd}\right] - \frac{1}{\rho}\chi\theta\sum_{i=1}^{n-1}\left\{\lambda_i\left[\underline{\sigma}_{iq}^{-1}\right]\right\}$$

This is greater then zero when:

$$\lambda_n\left[K - (\chi(1-\theta))K_{nd}\right] > \chi\theta\sum_{i=1}^{n-1}\left\{\lambda_i\left[\underline{\sigma}_{iq}^{-1}\right]\right\} \tag{51}$$

The above condition is always verified as long as the share of skilled investors in the market is sufficiently low:

$$\chi < \frac{\lambda_n K}{\theta\left\{\sum_{i=1}^{n-1}\left[\lambda_i\underline{\sigma}_{iq}^{-1}\right] - \lambda_l K_{nd}\right\} + \lambda_l K_{nd}} < 1$$

- When choosing to learn about an idiosyncratic shock l $K_{ld} = K$ (the condition is similar if sharing the capacity among a basket of idiosyncratic factors):

$$E[R_D - R_M] = \frac{1}{\rho}\lambda_l K - (\chi(1-\theta))\lambda_l K_{ld} - \frac{1}{\rho}\sum_{i=1}^{n-1}\left\{\lambda_i\left[\underline{\sigma}_{iq}^{-1}\chi\theta\right]\right\}$$

This is greater than zero when:

$$\lambda_l\left(K - (\chi(1-\theta))K_{ld} - \chi\theta\underline{\sigma}_{lq}^{-1}\right) > \chi\theta\sum_{i\neq j}^{n-1}\left\{\lambda_i\left[\underline{\sigma}_{iq}^{-1}\right]\right\} \text{ implies } \lambda_l\left(K - \chi K_{ld} - \chi\theta G_l\right) > \chi\theta\sum_{i\neq j}^{n-1}\left\{\lambda_i\left[\underline{\sigma}_{iq}^{-1}\right]\right\} \tag{52}$$

The share of skilled investors in the market for which the above condition is always verified is given by:

$$\chi < \frac{\lambda_l K}{\theta\left\{\sum_{i\neq j}^{n-1}\left[\lambda_i\underline{\sigma}_{iq}^{-1}\right] - \lambda_l G_l\right\} + \lambda_l K_{ld}} < 1$$

Effect of increases in σ_i:

Considering equation 50, when σ_i rises the only thing that is affected is λ_i, what needs to be taken into consideration is whether any attention was paid to that particular shock and whether the increase in σ_i causes any chance in attention.

- σ_n rises: if no attention was paid to the aggregate shock, the expected excess return of the discretionary skilled investors would be given by:

$$E[R_D - R_M] = \frac{1}{\rho}\sum_{j=1}^{J}\left\{\lambda_j\left[\sigma_{jd}^{-1} - (\chi(1-\theta))K_{jd} - \underline{\sigma}_{jq}^{-1}\chi\theta\right]\right\} - \frac{1}{\rho}\sum_{i\neq j}^{n-1}\left\{\lambda_i\left[\underline{\sigma}_{iq}^{-1}\chi\theta\right]\right\} \tag{53}$$

λ_n is continuous in σ_n so a marginal change in σ_n would not be enough to change the rankings in λs hence a marginal increase in σ_n would have no impact on the excess return. If the increase was a regime switch from expansions to recessions, instead, we would observe that the significant rise in σ_n would cause a change in the ranking of the λ_s and a likely shift in attention to the aggregate shock moving the expected excess return to:

$$E[R_D - R_M] = \frac{1}{\rho}\lambda_n K - (\chi(1-\theta))K_{nd} - \frac{1}{\rho}\sum_{i=1}^{n-1}\left\{\lambda_i\left[\underline{\sigma}_{iq}^{-1}\chi\theta\right]\right\} \tag{54}$$

If attention was already paid to the aggregate shock, hence the expected excess return was already as described in equation 54 above, a marginal increase in σ_n would determine an increase in λ_n and consequently an increase in the expected excess return. Hence overall an increase in σ_n weakly increases the expected excess return of discretionary skilled investors.

- $\sigma_{i\neq n\neq j}$ rises: if no attention was paid to idiosyncratic shocks, the expected excess return of the discretionary skilled investors would be as described in equation 54. Hence a marginal increase in the volatility of any of the idiosyncratic shocks would only determine an increase in λ_i and consequently a decrease in the expected excess return of the discretionary skilled investors.

- σ_j rises: in the case described in equation 53 discretionary skilled investors are optimally paying attention to a basket of j idiosyncratic shocks. When the volatility of one of these shocks rises, λ_j rises. This has a positive impact on the overall expected excess return if $\sigma_{jd}^{-1} - (\chi(1-\theta))K_{jd} - \underline{\sigma}_{jq}^{-1}\chi\theta > 0$ or equivalently if:

$$\theta < \frac{\sigma_{jd}^{-1} - \chi K_{jd}}{\chi G_j}$$

Effect of increases in θ:

$$\frac{\partial E[R_D - R_M]}{\partial \theta} = \frac{1}{\rho}\sum_{i=1}^{n-1}\left[\lambda_i \chi K_{iq}\left(e_{\lambda_i,\theta} - 1\right)\right] + \frac{1}{\rho}\sum_{j=1}^{J}\left[\lambda_j K_{jd}\left(\chi - \frac{1}{\theta}e_{\lambda_j,\theta}\left(1 - \chi + \chi\theta\right)\right)\right] + \frac{1}{\rho}\lambda_n K_{nd}\left[\frac{1}{\theta}e_{\lambda_n,\theta}\left(1 - \chi + \chi\theta\right) + \chi\right]$$

- When no attention is paid to the aggregate shock ($K_{nd} = 0$; $K_{jd} > 0$):

$$\frac{\partial E[R_D - R_M]}{\partial \theta} = \frac{1}{\rho}\sum_{i=1}^{n-1}\left[\lambda_i \chi K_{iq}\left(e_{\lambda_i,\theta} - 1\right)\right] + \frac{1}{\rho}\sum_{j=1}^{J}\left[\lambda_j K_{jd}\left(\chi - \frac{1}{\theta}e_{\lambda_j,\theta}\left(1 - \chi + \chi\theta\right)\right)\right]$$

  If assuming that the elasticities of the $\lambda_i s$ with respect to θ are less than 1; as θ increases the expected excess return of the discretionary investors decreases.

- When all attention is paid to the aggregate shock ($K_{nd} = K$; $K_{jd} = 0$):

$$\frac{\partial E[R_D - R_M]}{\partial \theta} = \frac{1}{\rho}\sum_{i=1}^{n-1}\left[\lambda_i \chi K_{iq}\left(e_{\lambda_i,\theta} - 1\right)\right] + \frac{1}{\rho}\lambda_n K_{nd}\left[\frac{1}{\theta}e_{\lambda_n,\theta}\left(1 - \chi + \chi\theta\right) + \chi\right] \quad (55)$$

  We have two counteracting effects as an increase in θ decreases the first part of equation 55 above but increases the second part.

Figure 1: **Number of quantitative and discretionary funds:** Number of Quantitative and Discretionary funds by month. The classification of funds into quantitative and discretionary is obtained using the Random Forest algorithm through the analysis of fund prospectuses.

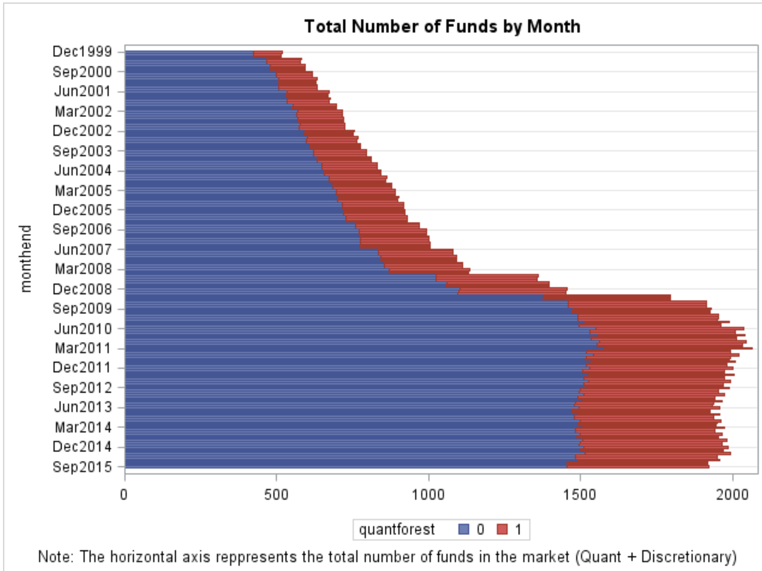


Figure 2: **Random Forest features importance:** Informativeness of the first 10 features used by the Random Forest algorithm. Informativeness is intended as the reduction in classification impurity (measured with entropy)

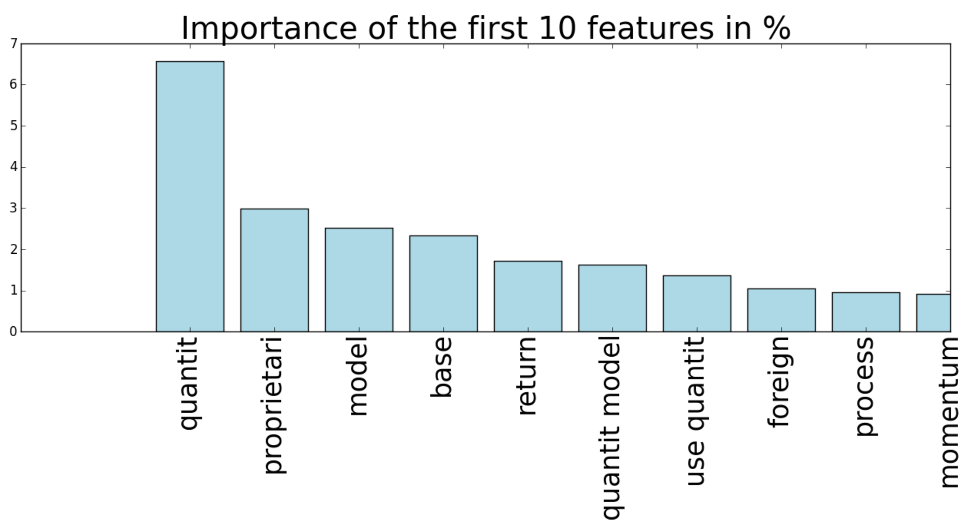
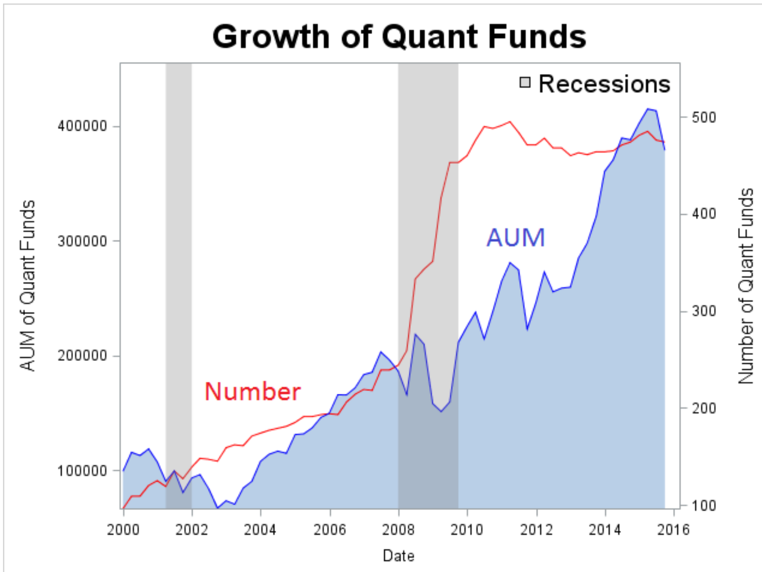


Figure 3: **Growth of quantitative funds:** Panel 1 shows the size of quantitative funds in terms of assets managed and number of funds. Panel 2 shows the relative size of quantitative funds in terms of assets managed and number of funds, as a percentage of overall market TNA and total number of funds respectively. The shaded areas indicate NBER recessions dates.



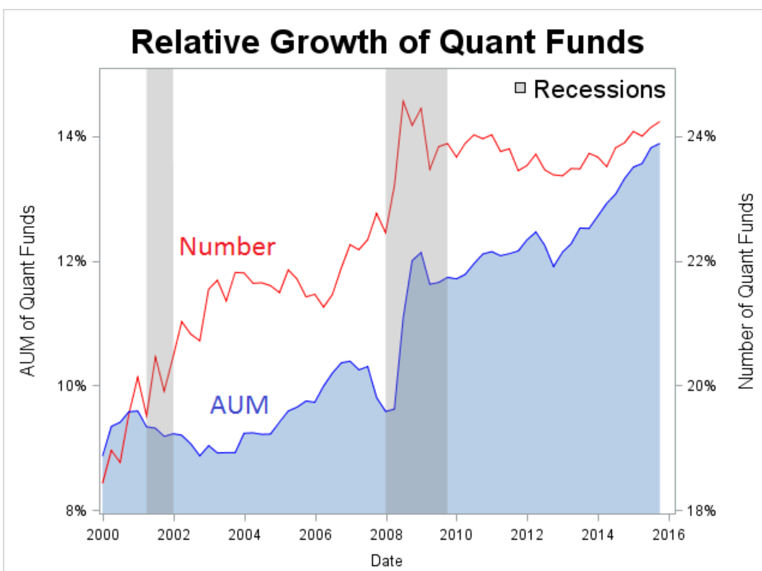


Figure 4: **Age of quantitative and discretionary funds:** Funds age by group over time. Mean age of quantitative and discretionary funds per month, for the period of 1999-2015. Age is measure in months from the fund's inception date.

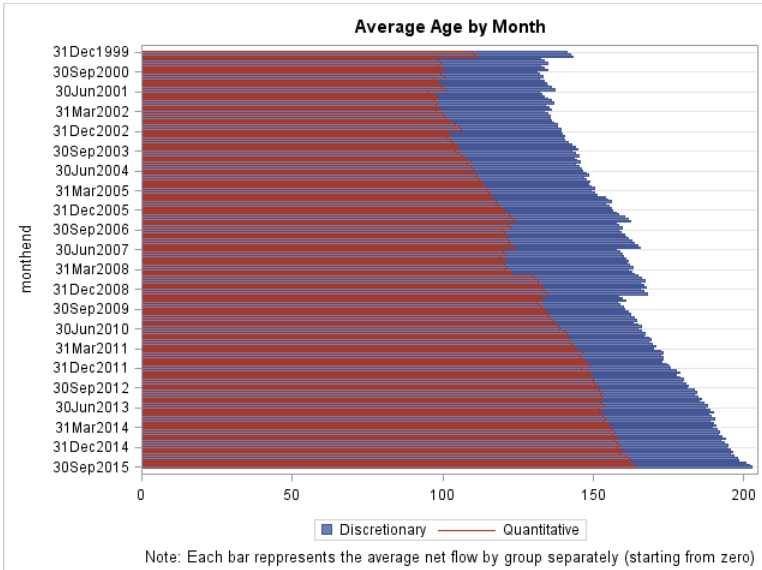


Figure 5: **Creation and Failure rates of quantitative and discretionary funds:** The creation and failure rates of quantitative (discretionary) funds are calculated as the number of quantitative (discretionary) funds who were born or died in a given month in percentage to the total number of quantitative (discretionary) funds present in the market.

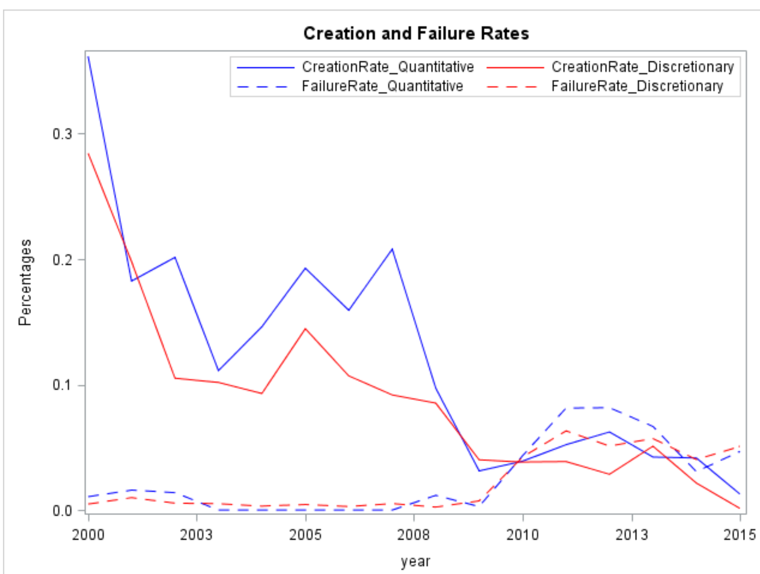


Figure 6: **Funds size by group:** Distribution of quantitative and discretionary log-fund-size across the entire sample.

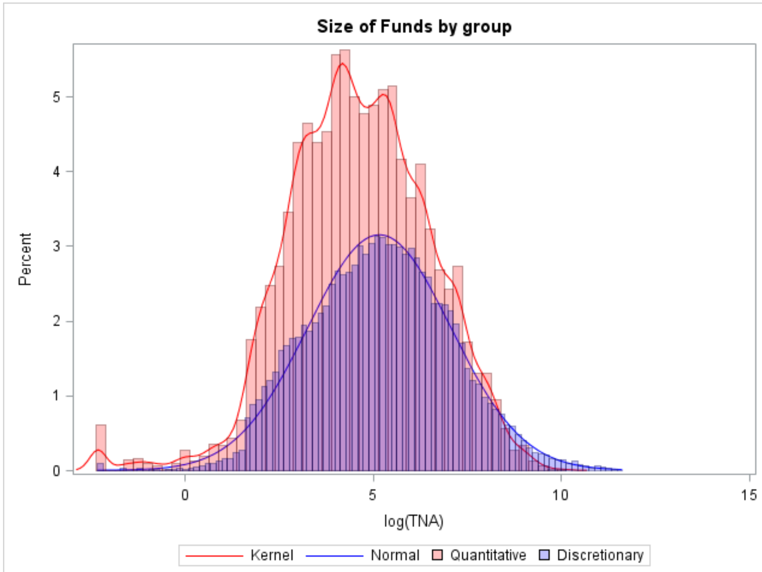


Figure 7: **Shortselling by quantitative and discretionary funds:** Average percentage of gross TNA sold short by discretionary and quantitative funds over time. The table represents the short positions over time of a sample of 50 funds (36 discretionary and 24 quantitative) for which short positions are available in the CRSP Mutual Fund Holdings dataset. The % of gross TNA sold short is computed as the relative value of the shares sold short relative to the gross TNA (value of long + short positions). Eg: for a 130-30 strategy the weight displayed in the table would be 30% of the fund's TNA divided by 160% of the fund's TNA.

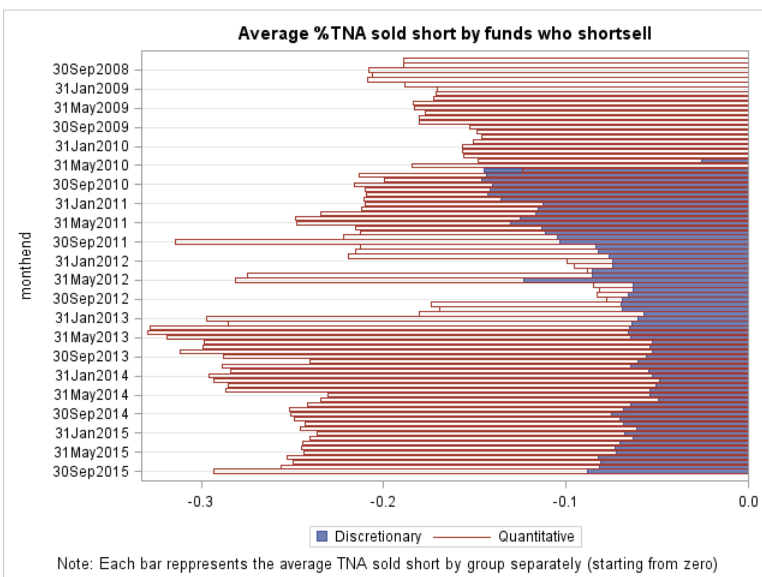


Figure 8: **Percentile difference in size of quantitative and discretionary funds:** The percentage size difference is estimated through a quantile regression of fund size ($log(TNA)$) on the $NBER_t$ and $Quant_j$ dummies and various de-trended fund level control variables (size, age, expenses ratio, turnover ratio, fund loads, net fund flows, fund flow volatility, as well as the load on value, momentum and size factors of their holdings – a detailed explanation of the variables construction can be found in the heading of Table 1): $TNA_{jt} = \alpha^p + \beta_1^p NBER_t + \beta_2^p Quant_j + \beta_3^p Quant_j \times NBER_t + \gamma^p X_{jt} + \varepsilon_t$. Standard errors are obtained through block bootstrapping (200 repetitions), which accounts for cross-sectional dependence across funds. Detrending allows to interpret α^p as the size of discretionary funds and β_2^p as the difference in size between quantitative and discretionary funds, for each percentile (identified by the superscript "p"). The observations in the graph represent the percentage difference in size of quantitative and discretionary funds: β_2^p/α^p, computed every 5 percentiles in addition to the 1^{st} and 99^{th} percentiles. The ratio is significantly different for most percentile combinations at the 1% level, with the exception of a few percentiles in the mid-section of the size distribution (i.e the 20^{th} percentile is not statistically different from the 25^{th} and 30^{th} and the 35^{th}, the 35^{th} percentile is not statistically different from the 40^{th} and the 85^{th} percentile is not statistically different from the 90^{th}).

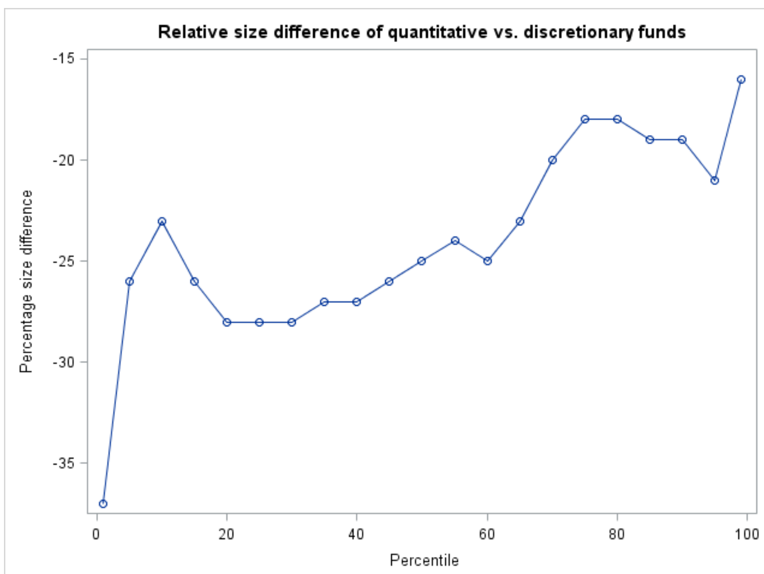
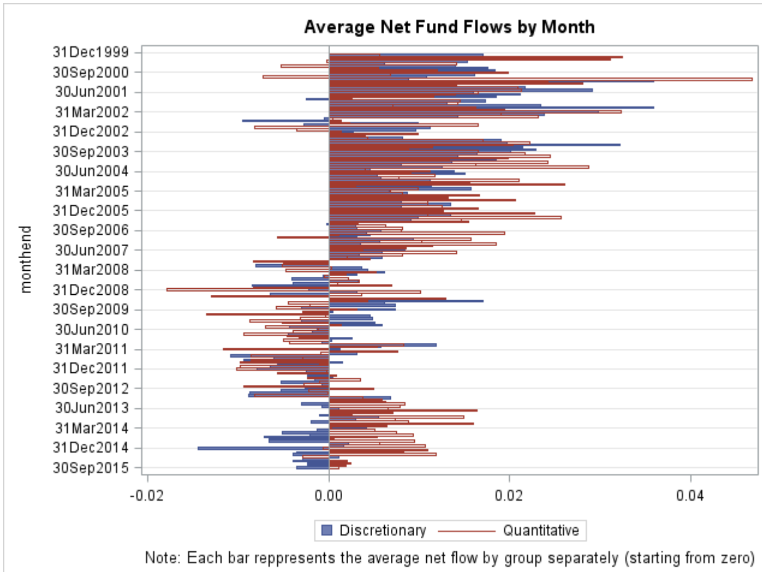


Figure 9: **Net fund flows of quantitative and discretionary funds:** Mean and median fund flows by group over time. Fund flows are measured as the monthly percentage change in TNA not due to returns.



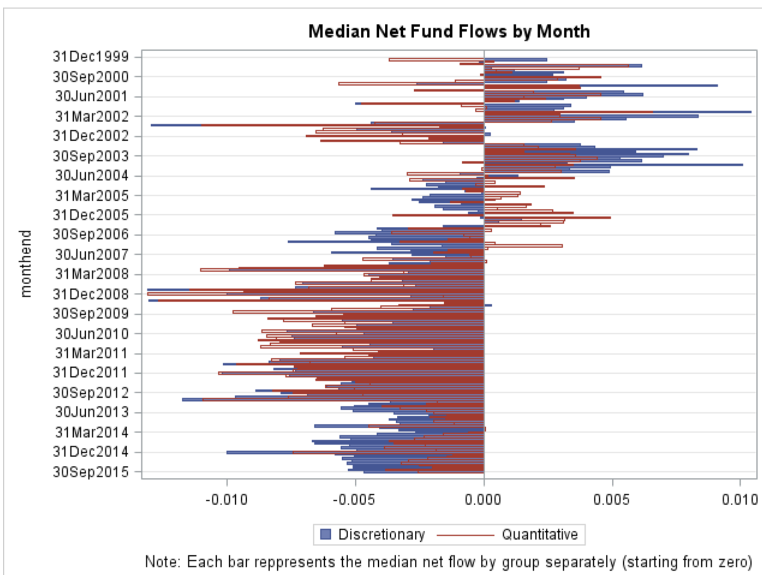


Figure 10: **Number of stocks held by quantitative and discretionary funds:** Distribution of the number of stocks held by quantitative and discretionary funds over the entire period. Panel 1: displays the distribution in expansions. Panel 2: display the distributions in recessions.

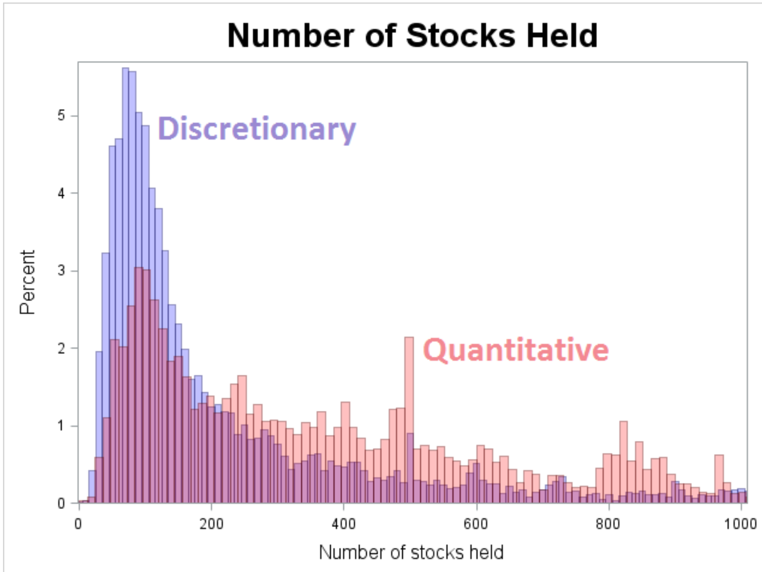


Figure 11: **Distribution of holdings dispersion and commonality of quantitative and discretionary funds:** Panel 1 displays the distribution of the commonality measure by group. Commonality is measured as the weighted average of the percentage of other quantitative (discretionary) funds who hold the same stocks. Weights are the normalized percentage of TNA that the fund invests in a certain stock (in excess of the market weight). Panel 2 displays the dispersion distribution by groups. Dispersion is measured as the cumulative squared weight weight allocated by funds to each stock in excess of the average weight allocated by all funds of their type.

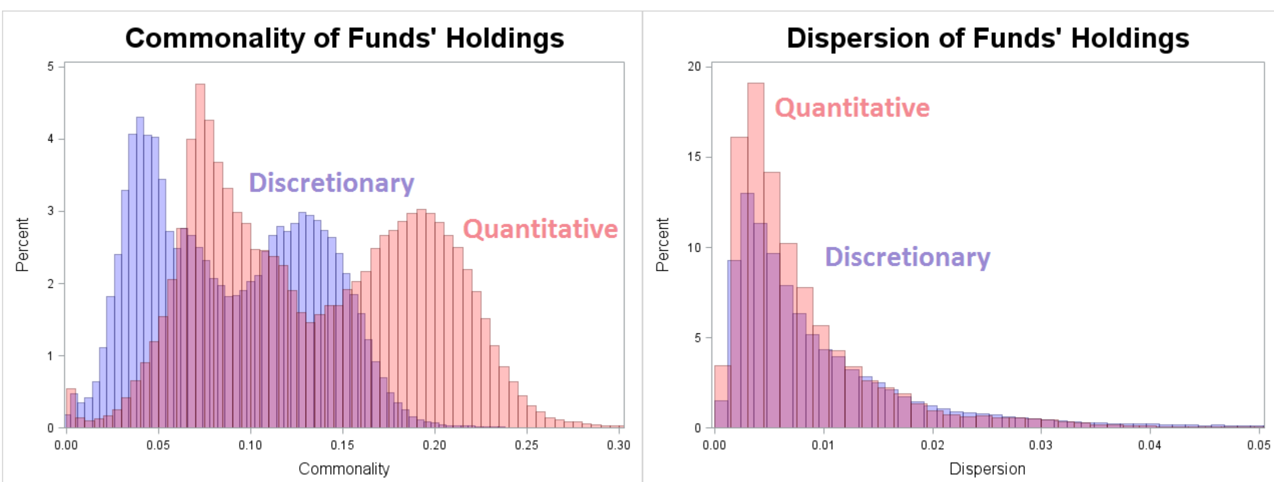


Table 1: **Differences in Age and Size of Quantitative and Discretionary Funds**
Dependent variables: funds age measured in months from inception (*Age*) and fund size measured with Total Net Assets (TNA). Independent variables: a dummy identifying recession periods (*Recession*), a dummy identifying quantitative funds (*Quant*), the interaction term between the *Quant* and *Recessions* dummies (*QuantXRecessions*), fund size measured as the natural logarithm of TNA ($Log(TNA)$), the winsorized funds expenses ratio (*ExpenseRatio*), the winsorized funds turnover ratio (*TurnoverRatio*), net fund flows measured as the monthly percentage change in TNA which is not determined by fund returns (*NetFlows*), total fund loads (*Load*), the 12–months rolling volatility of net fund flows (*FlowVolatility*), the natural logarithm of funds age in months from inception ($Log(Age)$) and the winsorized %TNA–weighted average load of the stocks held by the funds on the SMB (Small–Minus–Big – Size), HML (High–Minus–Low – Book–to–Market) and UMD (Momentum) factors (*Size*, *Value*, *Momentum* respectively). Regressions are run monthly and standard errors are clustered at the fund and month level.

	Age	*TNA*
	(1)	(2)
Constant	174.3***	1219.9***
	(0.000)	(0.000)
Recessions	-7.162**	20.84
	(0.025)	(0.828)
Quant	-18.51***	-667.0***
	(0.009)	(0.000)
QuantXRecessions	4.196*	21.50
	(0.083)	(0.751)
Log(TNA)	24.43***	
	(0.000)	
ExpenseRatio	-538.9	-173261.7***
	(0.517)	(0.000)
TurnoverRatio	-3.828	-249.2***
	(0.323)	(0.000)
NetFlows	-179.2***	2478.3***
	(0.000)	(0.000)
Loads	-1272.3**	15170.4
	(0.041)	(0.213)
FlowVolatility	-539.3***	-5520.7***
	(0.000)	(0.000)
Momentum	-1.148	302.4*
	(0.765)	(0.051)
Size	-24.24***	-278.0***
	(0.000)	(0.002)
Value	-14.22***	163.8
	(0.001)	(0.117)
Log(Age)		950.0***
		(0.000)
Observations	195417	195417

p-values in parentheses: * $p<0.10$, ** $p<0.05$, *** $p<0.01$

Table 2: **Differences in the Holdings of Quantitative and Discretionary Funds**
Dependent variables: the winsorized %TNA–weighted average load of the stocks held by each funds each month on the SMB (Small–Minus–Big – Size), HML (High–Minus–Low – Book–to–Market) and UMD (Momentum) factors (*Size*, *Value*, *Momentum* respectively), the winsorized turnover ratio in fund holdings (*TurnoverRatio*) and the amount of cash held (*Cash*). Independent variables are the same as described in Table 1. Regressions are run monthly and standard errors are clustered at the fund and month level.

	Momentum	*Size*	*Value*	*TurnoverRatio*	*Cash*
	(1)	(2)	(3)	(4)	(5)
Constant	0.00461	0.272***	-0.0594***	0.757***	3.174***
	(0.660)	(0.000)	(0.000)	(0.000)	(0.000)
Recessions	-0.0107	-0.0625**	0.0753*	0.0976***	0.504***
	(0.580)	(0.023)	(0.074)	(0.000)	(0.000)
Quant	0.00822*	-0.0276	0.0770***	0.147***	-0.730***
	(0.051)	(0.210)	(0.000)	(0.000)	(0.000)
QuantXRecessions	-0.00977*	0.0314**	-0.00492	0.00737	-0.668**
	(0.098)	(0.016)	(0.781)	(0.793)	(0.021)
Log(TNA)	0.00370***	0.00745*	0.00268	-0.0175***	0.145***
	(0.000)	(0.068)	(0.411)	(0.007)	(0.000)
ExpenseRatio	2.093***	24.39***	1.414	23.05***	125.4***
	(0.000)	(0.000)	(0.399)	(0.000)	(0.000)
TurnoverRatio	0.0363***	0.0356***	-0.0662***		-0.274**
	(0.000)	(0.001)	(0.000)		(0.034)
NetFlows	0.0884***	0.0436	0.289***	-0.355***	4.225***
	(0.005)	(0.335)	(0.000)	(0.000)	(0.000)
Loads	0.936**		9.626***	-6.570	125.8
	(0.029)		(0.001)	(0.220)	(0.171)
FlowVolatility	0.0102	0.517***	0.456***	1.695***	3.325**
	(0.718)	(0.000)	(0.000)	(0.000)	(0.046)
Size	-0.0141		0.0946***	0.0714***	0.541***
	(0.479)		(0.000)	(0.002)	(0.000)
Value	-0.0194	0.0979***		-0.144***	0.517***
	(0.401)	(0.000)		(0.000)	(0.000)
Loads		0.480			
		(0.894)			
Momentum		-0.0327	-0.0433	0.174***	-0.500***
		(0.491)	(0.419)	(0.000)	(0.004)
Log(Age)				-0.0295**	-0.333***
				(0.048)	(0.000)
Observations	195417	195417	195417	195417	173924

p-values in parentheses: * $p<0.10$, ** $p<0.05$, *** $p<0.01$

Table 3: **Differences in NetFundFlows of Quantitative and Discretionary Funds**

	$NetFlows$	$FlowVolatility$	$NetFlows2011$
	(1)	(2)	(3)
$Constant$	0.00175***	0.0337***	-0.00183**
	(0.001)	(0.000)	(0.012)
$Recessions$	-0.000606	-0.000691	
	(0.704)	(0.313)	
$Quant$	-0.00104	0.00135	0.00449***
	(0.191)	(0.252)	(0.000)
$QuantXRecessions$	-0.00520***	0.00269	
	(0.005)	(0.101)	
$Log(Age)$	-0.00999***	-0.00891***	
	(0.000)	(0.000)	
$Log(TNA)$	0.00321***	-0.00392***	
	(0.000)	(0.000)	
$ExpenseRatio$	0.390***	-0.496***	
	(0.000)	(0.000)	
$TurnoverRatio$	-0.00252***	0.00487***	
	(0.000)	(0.000)	
$Loads$	0.177	-0.0407	
	(0.271)	(0.766)	
$FlowVolatility$	0.329***		
	(0.000)		
$Momentum$	0.00324***	0.000341	
	(0.005)	(0.677)	
$Size$	0.0000805	0.00254***	
	(0.901)	(0.001)	
$Value$	0.00322***	0.00172**	
	(0.000)	(0.014)	
$NetFlows$		0.132***	
		(0.000)	
$Observations$	195417	195417	89772

p-values in parentheses: * p<0.10, ** p<0.05, *** p<0.01

Table 4: **Differences in Fees Charged by Quantitative and Discretionary Funds**
Dependent variables: funds expenses ratio (*ExpenseRatio*), management fees (*ManagementFees*), Actual_12b1 fee (*Actual_12b1*) and fund loads (*Loads*). Independent variables are the same as described in Table 1. Regressions are run monthly and standard errors are clustered at the fund and month level.

	ExpenseRatio	*ManagementFees*	*Actual_12b1*	*Loads*
	(1)	(2)	(3)	(4)
Constant	0.0121***	0.00694***	0.00365***	0.000233***
	(0.000)	(0.000)	(0.000)	(0.000)
Recessions	0.0000578	0.000489***	0.0000662	-0.0000446*
	(0.594)	(0.000)	(0.299)	(0.067)
Quant	-0.00135***	-0.000625***	0.0000556	-0.00000609
	(0.000)	(0.000)	(0.706)	(0.922)
QuantXRecessions	-0.000152	-0.000152	-0.000155**	0.0000492
	(0.137)	(0.303)	(0.043)	(0.325)
Log(Age)	0.0000247	0.000545***	-0.000171**	-0.0000302
	(0.790)	(0.000)	(0.016)	(0.205)
Log(TNA)	-0.000673***	0.000199***	-0.000000974	0.0000236
	(0.000)	(0.000)	(0.972)	(0.147)
TurnoverRatio	0.000674***	0.000427***	0.000325***	-0.0000463
	(0.000)	(0.000)	(0.000)	(0.309)
NetFlows	0.00159***	-0.00482***	-0.0000235	0.000173
	(0.000)	(0.000)	(0.930)	(0.395)
Loads	0.186***	0.128***	-0.0157	
	(0.000)	(0.000)	(0.196)	
FlowVolatility	-0.00504***	-0.00660***	-0.00314***	-0.0000994
	(0.000)	(0.000)	(0.009)	(0.769)
Momentum	0.000290***	0.000222**	0.0000401	0.0000319
	(0.008)	(0.048)	(0.553)	(0.523)
Size	0.00146***	0.00145***	-0.0000852	0.00000510
	(0.000)	(0.000)	(0.353)	(0.923)
Value	0.0000904	0.0000329	0.0000488	0.000140*
	(0.387)	(0.732)	(0.529)	(0.087)
ExpenseRatio				0.0447***
				(0.001)
Observations	195417	195389	126070	195417

p-values in parentheses: * p<0.10, ** p<0.05, *** p<0.01

Table 5: **Test of Model Assumptions:**
Discretionary Funds Shift Attention, Quantitative Funds Specialize

Dependent variables: funds' timing ability in recessions (*TimingRecession*) and expansions (*TimingExpansion*) and their picking ability in recessions (*PickingRecession*). Market timing and stock picking are measured as the covariance of funds' allocations (in excess of the market weight) with future industrial production growth and earnings surprises respectively, as defined is Section 5.1. Independent variables: dummies identifying the top 25% of funds that most frequently display high stock picking ability in expansions (*TopPickersExpansions*) and high market timing ability in recessions (*TopTimersRecessions*) respectively and interaction terms between the *Quant* and *TopTimersRecessions* dummies (*QuantXTopTimersRecessions*) and between the *Quant* and *TopPickersExpansions* dummies (*QuantXTopPickersExpansions*), also defined is Section 5.1. All other independent variables are described in Table 1. Regressions are run monthly, standard errors are clustered at the fund and month level.

	TimingRecession	*PickingRecession*	*TimingExpansion*
	(1)	(2)	(3)
Constant	0.664***	-1.072***	-0.203***
	(0.000)	(0.000)	(0.000)
Quant	-0.196***	0.356***	0.0405***
	(0.000)	(0.000)	(0.001)
TopPickersExpansions	0.454***	-0.864***	
	(0.000)	(0.000)	
QuantXTopPickersExpansions	-0.621***	0.248***	
	(0.000)	(0.004)	
TopTimersRecessions			0.0323
			(0.494)
QuantXTopTimersRecessions			-0.0412
			(0.352)
Log(Age)	-0.0402	-0.0607*	-0.0512***
	(0.124)	(0.059)	(0.000)
Log(TNA)	-0.00613	0.0923***	0.0226***
	(0.609)	(0.000)	(0.000)
ExpenseRatio	32.31***	-19.85**	-3.768*
	(0.000)	(0.040)	(0.063)
TurnoverRatio	-0.285***	0.413***	-0.0491***
	(0.000)	(0.000)	(0.000)
NetFlows	-0.739**	1.936***	-0.249*
	(0.014)	(0.000)	(0.072)
Loads	-32.49***	38.38***	2.582
	(0.000)	(0.000)	(0.127)
FlowVolatility	0.337	-2.318***	0.0351
	(0.225)	(0.000)	(0.715)
Momentum	-1.345***	1.493***	0.204***
	(0.001)	(0.009)	(0.003)
Size	0.381***	0.210*	0.136***
	(0.000)	(0.076)	(0.000)
Value	0.843***	-0.360***	0.162***
	(0.000)	(0.004)	(0.001)
Observations	24625	24670	152869

p-values in parentheses: * $p<0.10$, ** $p<0.05$, *** $p<0.01$

Table 6: **Number of Stocks Held by Quantitative and Discretionary Funds**
Dependent variable: number of stocks held (*Number_holdings*). Independent variables: dummies for quantitative funds (*Quant*) and recessions (*Recessions*) and their interaction (*QuantXRecessions*); amount of cash held (*Cash*) and the fund log–size measured as the natural logarithm of total net assets (*Log*(*TNA*)).

	Number of Stocks Held	
	(1)	(2)
Constant	117.6***	160.2***
	(0.000)	(0.000)
Quant	108.4***	124.8***
	(0.000)	(0.000)
Recessions	-4.985	2.067
	(0.128)	(0.648)
QuantXRecessions	21.27***	12.85
	(0.005)	(0.128)
Cash		-0.722
		(0.274)
Log(*TNA*)		26.47***
		(0.000)
Observations	229742	173926

p-values in parentheses: * p<0.10, ** p<0.05, *** p<0.01

Table 7: **Portfolio Risk of Quantitative and Discretionary Funds**

Dependent variables: 12–months and 36–months rolling volatility of fund returns (Vol_12m and Vol_36m respectively), CAPM idiosyncratic volatility ($IVCapm\_12m$ and $IVCapm\_36m$ respectively), Fama–French 3–factors idiosyncratic volatility ($IVFF\_12m$ and $IVFF\_36m$ respectively) and CAR 4–factor idiosyncratic volatility ($IVCar\_12m$ and $IVCar\_36m$ respectively). Idiosyncratic volatility is calculated as the volatility of the residual of a regression of fund returns on the market factor (CAPM) the size and value factors (3–factors) and the momentum factor (4–factors). Independent variables are the same as described in Table 1. Regressions are run monthly, standard errors are clustered at the fund and month level.

	Vol_12m	Vol_36m	$IVolCapm\_12m$	$IVolCapm\_36m$	$IVolFF\_12m$	$IVolFF\_36m$	$IVolCar\_12m$	$IVolCar\_36m$
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
$Constant$	4.319***	4.915***	1.578***	1.763***	1.192***	1.413***	1.150***	1.365***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
$Recessions$	1.780***		0.706***		0.740***		0.697***	
	(0.000)		(0.000)		(0.000)		(0.000)	
$Quant$	-0.0825**	-0.108***	-0.152***	-0.147***	-0.183***	-0.186***	-0.179***	-0.186***
	(0.012)	(0.008)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
$QuantXRecess$	-0.146***		-0.0616*		-0.0385		-0.0351	
	(0.002)		(0.056)		(0.186)		(0.205)	
$Log(Age)$	-0.0274	0.00279	-0.0735***	-0.0647***	-0.0569***	-0.0430***	-0.0482***	-0.0315**
	(0.405)	(0.936)	(0.000)	(0.000)	(0.000)	(0.003)	(0.000)	(0.019)
$Log(TNA)$	-0.0170	-0.00316	0.0322***	0.0319***	0.0236***	0.0229***	0.0202***	0.0196***
	(0.258)	(0.831)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.001)
$ExpenseRatio$	16.03**	30.46***	44.43***	53.94***	44.82***	51.33***	42.80***	48.94***
	(0.023)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
$TurnoverRatio$	0.336***	0.414***	0.212***	0.261***	0.195***	0.244***	0.172***	0.213***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
$NetFlows$	-0.886***	-0.737***	-0.109	-0.0853	-0.00832	-0.0622	-0.00967	-0.0714
	(0.000)	(0.001)	(0.323)	(0.451)	(0.921)	(0.498)	(0.901)	(0.396)
$Loads$	7.541	4.323	6.141	6.092	1.612	2.809	0.287	1.927
	(0.216)	(0.579)	(0.178)	(0.210)	(0.672)	(0.524)	(0.937)	(0.647)
$FlowVolatility$	1.472***	0.494	1.927***	1.025***	1.443***	0.865***	1.417***	0.928***
	(0.000)	(0.250)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
$Momentum$	-0.421***	-0.690***	0.0445	0.0173	-0.110**	-0.0778*	-0.144***	-0.0926**
	(0.010)	(0.000)	(0.461)	(0.746)	(0.014)	(0.065)	(0.000)	(0.025)
$Size$	0.854***	0.688***	0.791***	0.724***	0.239***	0.272***	0.266***	0.283***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
$Value$	-0.436***	-0.213***	0.00469	-0.00253	-0.105***	-0.0652**	-0.107***	-0.0745***
	(0.000)	(0.001)	(0.914)	(0.949)	(0.001)	(0.017)	(0.001)	(0.004)
$Observations$	186427	186427	187822	178979	187822	178979	187822	178979

p-values in parentheses: * p<0.10, ** p<0.05, *** p<0.01

Table 8: **Market Timing and Stock Picking Ability of Discretionary Funds**
Dependent variables: picking and timing abilities of discretionary investors (*Picking*, *Timing*), measured as the covariance of portfolio weights (in excess of the market weight) with future earning surprises and the growth in industrial production respectively, as defined is Section 5.1. Independent variables: a dummy variable identifying recession periods (*Recessions*) , the detrended share of assets managed by quantitative funds with respect to the assets managed by all funds in my sample (*Theta_TNA*)–detrending is obtained with Hodrick-Prescott filtering–and the percentage of the total US stock market capitalization held by quantitative funds in my sample (*Theta_Hold*). All other independent variables are described in Table 1. Regressions are run monthly, standard errors are clustered at the fund and month level.

	Discretionary			
	Picking	*Timing*	*Picking*	*Timing*
	(1)	(2)	(3)	(4)
Constant	-0.0206	0.121	-0.210***	0.798***
	(0.790)	(0.209)	(0.004)	(0.000)
Recessions	-1.567***	1.030***	-1.468***	0.962***
	(0.000)	(0.000)	(0.000)	(0.000)
Theta_TNA	-0.215***	0.175***		
	(0.000)	(0.006)		
Theta_Hold			-0.658***	0.151***
			(0.000)	(0.003)
Log(Age)	0.00499	-0.0741***	0.0142	-0.0522***
	(0.700)	(0.000)	(0.150)	(0.000)
Log(TNA)	0.00618	0.0169**	0.00193	0.00712
	(0.388)	(0.040)	(0.679)	(0.159)
ExpenseRatio	2.202	3.093	2.181	-2.857
	(0.484)	(0.380)	(0.351)	(0.285)
TurnoverRatio	0.194***	-0.122***	0.187***	-0.137***
	(0.000)	(0.000)	(0.000)	(0.000)
NetFlows	0.805***	-0.453***	0.763***	-0.515***
	(0.000)	(0.003)	(0.000)	(0.000)
Loads	4.931*	-4.843	3.148	-4.646*
	(0.053)	(0.129)	(0.207)	(0.089)
FlowVolatility	-0.216	-0.0750	-0.205	0.0106
	(0.415)	(0.774)	(0.412)	(0.967)
Momentum	-0.0489	0.124*	-0.0582*	0.125***
	(0.516)	(0.089)	(0.056)	(0.000)
Size	0.0276	0.176***	0.0359*	0.178***
	(0.380)	(0.000)	(0.054)	(0.000)
Value	-0.0254	0.204***	-0.0914***	0.217***
	(0.451)	(0.000)	(0.000)	(0.000)
DiscretionaryHold			0.116***	-0.111***
			(0.000)	(0.000)
Observations	149345	143195	149345	143195

p-values in parentheses: * $p<0.10$, ** $p<0.05$, *** $p<0.01$

Table 9: **Size of Stocks Held by Discretionary Funds**
Independent variable: average market capitalization of stocks stocks held measured with simple average (*MarketValue*), average weighted by %TNA (*MarketValue_w*) and weighted by %TNA in excess of the market weight (*MarketValue_wd*), as defined is Section 5.1. Dependent variables: dummy variables identifying discretionary funds (*Discretionary*), and recessions (*Recessions*) and an interaction term between the two dummies (*DiscretXRecessions*).

	MarketValue	*MarketValue_w*	*MarketValue_wd*
	(1)	(2)	(3)
Constant	24449575.8***	42867071.0***	33420522.1***
	(0.000)	(0.000)	(0.000)
Discretionary	599454.1	-7149830.3**	-5664176.9*
	(0.697)	(0.015)	(0.080)
Recessions	-1055761.5	-2079402.3	-2243909.5
	(0.306)	(0.280)	(0.439)
DiscretXRecessions	-149161.7	666792.0	589729.2
	(0.834)	(0.712)	(0.836)
Observations	133957	133957	133957

p-values in parentheses: * p<0.10, ** p<0.05, *** p<0.01

Table 10: **Age of Stocks Held by Discretionary Funds**
Dependent variable: average age of stocks held (in months from inception), simple average (*Stocks_Age*), %TNA–weighted average (*Stocks_Age_w*) and %TNA–difference weighted average (*Stocks_Age_wd*), as defined is Section 5.1. Dependent variables as in table 9.

	Stocks_Age	*Stocks_Age_w*		*Stocks_Age_wd*	
	(1)	(2)	(3)	(4)	(5)
Constant	317.0***	340.3***	335.8***	335.0***	327.6***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Discretionary	-27.84***	-35.51***	-34.77***	-36.81***	-35.58***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Recessions	-5.815	5.820	6.036	3.051	3.410
	(0.239)	(0.370)	(0.351)	(0.606)	(0.569)
DiscretXRecessions	0.510	-4.968	-5.038	-2.301	-2.416
	(0.809)	(0.140)	(0.135)	(0.387)	(0.358)
MarketValue_w			0.000000104		0.000000172
			(0.570)		(0.274)
Observations	133957	133957	133957	133957	133957

p-values in parentheses: * p<0.10, ** p<0.05, *** p<0.01

Table 11: **Media Mentions of Stocks Held by Discretionary Funds**
Dependent variable: average number of times stock held are mentioned in Dow Jones news per month, simple average (*Media*), %TNA–weighted average (*Media_w*) and %TNA–difference weighted average (*Media_wd*), as defined is Section 5.1. Dependent variables as in table 9.

	Media	*Media_w*		*Media_wd*	
	(1)	(2)	(3)	(4)	(5)
Constant	245.5***	352.3***	-67.51***	293.8***	275.0***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Discretionary	8.622	-56.64**	13.37	-33.77**	-30.63**
	(0.527)	(0.014)	(0.175)	(0.035)	(0.043)
Recessions	19.57*	8.349	28.71*	21.65	22.56
	(0.087)	(0.692)	(0.069)	(0.134)	(0.113)
DiscretXRecessions	17.17***	20.82	14.29*	3.884	3.591
	(0.001)	(0.173)	(0.076)	(0.542)	(0.570)
MarketValue_w			0.00000979***		0.000000439
			(0.000)		(0.280)
Observations	133957	133957	133957	133957	133957

p-values in parentheses: * $p<0.10$, ** $p<0.05$, *** $p<0.01$

Table 12: **Analysts Coverage of Stocks Held by Discretionary Funds**
Dependent variable: average number of forecasts made by analysts on stocks held, simple average (*Analysts*), %TNA–weighted average (*Analysts_w*) and %TNA–difference weighted average (*Analysts_wd*), as defined is Section 5.1. Dependent variables as in table 9.

	Analysts	*Analysts_w*		*Analysts_wd*	
	(1)	(2)	(3)	(4)	(5)
Constant	374.7***	192.7***	410.1***	394.9***	385.6***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Discretionary	5.115	20.57**	-15.69	-6.431	-4.880
	(0.619)	(0.031)	(0.249)	(0.556)	(0.647)
Recessions	-121.9***	-131.2***	-141.7***	-130.3***	-129.8***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
DiscretXRecessions	-0.641	8.862	12.24	2.401	2.256
	(0.924)	(0.259)	(0.212)	(0.715)	(0.732)
MarketValue_w		0.00000507***			0.000000217
		(0.000)			(0.266)
Observations	133957	133957	133957	133957	133957

p-values in parentheses: * $p<0.10$, ** $p<0.05$, *** $p<0.01$

Table 13: **Performance of Funds Holding Stocks with Smaller Information Gap**
Dependent variables: 12–months rolling CAPM, Fama–French 3–factors and Carhart 4–factors alphas. Independent variables: dummy variables that identify recession periods (*Recessions*) and discretionary funds (*Discretionary*) and an interaction between the two (*DiscretXRecessions*); active information gap measured in terms of the number of analysts following the stocks held (*Analysts_wd*), the age of stocks held (*Stocks_Age_wd*) and the number of times stocks held are mentioned monthly in Dow Jones news (*Media_wd*). Are also included interaction terms between all active information gap measures and the discretionary dummy (*DisctretXAnalysts_wd*, *DisctretXStocks_Age_wd*, *DisctretXMedia_wd*). The active information gap is constructed as the weighted average of the age, number of media mentions and number of analysts forecasts for the stocks held each month by each fund. The weights used are the normalized weight differences of the % of TNA allocated by each fund to stocks minus their weight in the market, as defined is Section 5.1. Control variables are omitted for brevity; the controls used are the same in Table 1. Regressions are run monthly, standard errors are clustered at the fund and month level.

	12 Months Rolling Alphas		
	CAPM	$3-Factor$	$4-Factor$
	(1)	(2)	(3)
Recessions	0.126***	0.0814***	0.0164
	(0.000)	(0.000)	(0.132)
Discretionary	0.0771***	0.156***	0.119***
	(0.000)	(0.000)	(0.000)
DiscretXRecessions	0.0925***	0.0167	0.0213*
	(0.000)	(0.232)	(0.099)
Analysts_wd	-0.000328***	-0.000188***	-0.000261***
	(0.000)	(0.000)	(0.000)
DisctretXAnalysts_wd	-0.000117***	-0.000201***	-0.000163***
	(0.000)	(0.000)	(0.000)
Stocks_Age_wd	-0.000349***	0.000297***	-0.0000341
	(0.000)	(0.000)	(0.402)
DisctretXStocks_Age_wd	0.0000329	-0.000118**	-0.0000619
	(0.522)	(0.017)	(0.177)
Media_wd	-0.000313***	-0.000247***	-0.000228***
	(0.000)	(0.000)	(0.000)
DisctretXMedia_wd	-0.000120***	-0.0000458***	-0.0000567***
	(0.000)	(0.010)	(0.000)
Constant	0.895***	0.184***	0.360***
	(0.000)	(0.000)	(0.000)
Observations	127757	127757	127757

p-values in parentheses * p<0.10, ** p<0.05, *** p<0.01

Table 14: **Differences in Holdings Dispersion and Commonality among Discretionary and Quantitative funds**

Dependent variables: dispersion (*Dispersion*) and commonality (*Commonality*) in holdings. Whereby commonality is defined as the allocation-weighted average of the percentage of funds of the same group which hold each stock over time and Dispersion is defined as the cumulative squared difference in the weight allocated by each funds and the average weight allocated by funds of the same type to each stock, as described is Section 5.1. Independent variables: a dummy indicating discretionary funds (*Discretionary*), a dummy indicating recession periods (*Recessions*)and an interaction term between the two dummies (*DiscretXRecessions*). All other variables are described in Table 1. Regressions are run monthly, standard errors are clustered at the fund and month level.

	Dispersion	*Commonality*
	(1)	(2)
intercept	0.00839***	14.79***
	(0.000)	(0.000)
Discretionary	0.00257***	-5.782***
	(0.000)	(0.000)
Recessions	0.00173***	1.140***
	(0.005)	(0.000)
DiscretXRecessions	0.000135	-1.353***
	(0.783)	(0.000)
Controls	Yes	Yes
Clustering	Yes	Yes

p-values in parentheses * p<0.10, ** p<0.05, *** p<0.01

Table 15: **Performance of Quantitative and Discretionary Funds (CAPM)**
Independent variables: performance measured as the 12–months and 36–months rolling CAPM alphas (windsorized at the 1% level). Dependent variables as in table 9. Regressions are run monthly, standard errors are clustered at the fund and month level.

	CAPM Alphas			
	12 *Months Rolling*			36 *Months Rolling*
	(1)	(2)	(3)	(4)
Constant	-0.0328	-0.0266	-0.0286	0.0360*
	(0.131)	(0.264)	(0.209)	(0.069)
Recessions	0.282***	0.319***	0.310***	0.0479*
	(0.000)	(0.000)	(0.000)	(0.059)
Quant		-0.0268*	-0.0269**	-0.0214*
		(0.069)	(0.050)	(0.098)
QuantXRecessions		-0.164***	-0.153***	-0.0884***
		(0.000)	(0.000)	(0.001)
log(Age)			-0.0286***	-0.0454***
			(0.001)	(0.000)
Log(TNA)			0.0288***	0.0393***
			(0.000)	(0.000)
ExpenseRatio			8.080***	10.29***
			(0.000)	(0.000)
TurnoverRatio			-0.0200	-0.0296***
			(0.145)	(0.003)
NetFlows			1.952***	1.125***
			(0.000)	(0.000)
Loads			7.027**	6.208**
			(0.019)	(0.029)
FlowVolatility			1.098***	1.003***
			(0.000)	(0.000)
Momentum			0.0488	0.0440
			(0.302)	(0.188)
Size			0.0586	0.0852***
			(0.124)	(0.000)
Value			0.163***	0.134***
			(0.000)	(0.000)
Observations	199566	199566	187822	178979

p-values in parentheses: * p<0.10, ** p<0.05, *** p<0.01

Table 16: **Performance of Quantitative and Discretionary Funds (3-Factor Model)**
Independent variables: performance measured as the 12–months and 36–months rolling Fama–French 3–factor alphas (excess return on the market, SMB, HML – windsorized at the 1% level). Dependent variables as in table 9. Regressions are run monthly, standard errors are clustered at the fund and month level.

	$3-Factor\ Alphas$			
	12 $Months\ Rolling$			36 $Months\ Rolling$
	(1)	(2)	(3)	(4)
$Constant$	-0.0775***	-0.0710***	-0.0672***	-0.0329***
	(0.000)	(0.000)	(0.000)	(0.004)
recession	0.0641	0.0864**	0.0849**	0.107***
	(0.112)	(0.039)	(0.035)	(0.000)
$Quant$		-0.0284**	-0.0208*	-0.0303***
		(0.031)	(0.091)	(0.004)
$QuantXRecessions$		-0.0979***	-0.0886***	-0.0862***
		(0.001)	(0.002)	(0.000)
$Log(Age)$			-0.0130**	-0.0198***
			(0.047)	(0.001)
$Log(TNA)$			0.0183***	0.0276***
			(0.000)	(0.000)
$ExpenseRatio$			1.026	1.626
			(0.457)	(0.169)
$TurnoverRatio$			-0.0816***	-0.0660***
			(0.000)	(0.000)
$NetFlows$			1.417***	0.872***
			(0.000)	(0.000)
$Loads$			0.954	1.402
			(0.714)	(0.396)
$FlowVolatility$			0.826***	0.859***
			(0.000)	(0.000)
$Momentum$			0.0529	0.0608***
			(0.114)	(0.002)
$Size$			0.0370*	0.0289***
			(0.091)	(0.002)
$Value$			0.0649*	0.0111
			(0.065)	(0.446)
$Observations$	199566	199566	187822	178979

p-values in parentheses: * p<0.10, ** p<0.05, *** p<0.01

Table 17: **Performance of Quantitative and Discretionary Funds (4-Factor Model)**
Independent variables: performance measured as the 12–months and 36–months rolling CAR 4–factor alphas (excess return on the market, SMB, HML, momentum – windsorized at the 1% level). Dependent variables as in table 9. Regressions are run monthly, standard errors are clustered at the fund and month level.

	$4 - Factor\ Alphas$			
	$12\ Months\ Rolling$			$36\ Months\ Rolling$
	(1)	(2)	(3)	(4)
$Constant$	-0.0567***	-0.0506***	-0.0450***	-0.0272***
	(0.000)	(0.002)	(0.003)	(0.007)
$Recessions$	-0.0387	-0.0238	-0.0291	0.0426***
	(0.465)	(0.664)	(0.594)	(0.009)
$Quant$		-0.0265**	-0.0204*	-0.0275***
		(0.026)	(0.058)	(0.005)
$QuantXRecessions$		-0.0670**	-0.0604**	-0.0901***
		(0.018)	(0.032)	(0.000)
$Log(Age)$			-0.0174***	-0.0168***
			(0.008)	(0.004)
$Log(TNA)$			0.0188***	0.0249***
			(0.000)	(0.000)
$ExpenseRatio$			1.903	0.690
			(0.161)	(0.545)
$TurnoverRatio$			-0.0692***	-0.0624***
			(0.000)	(0.000)
$NetFlows$			1.296***	0.830***
			(0.000)	(0.000)
$Loads$			0.678	1.233
			(0.790)	(0.421)
$FlowVolatility$			0.740***	0.827***
			(0.000)	(0.000)
$Momentum$			-0.118**	0.0381**
			(0.030)	(0.017)
$Size$			-0.0243	0.0427***
			(0.368)	(0.000)
$Value$			0.0533	0.0170
			(0.278)	(0.195)
$Observations$	199566	199566	187822	178979

p-values in parentheses: * p<0.10, ** p<0.05, *** p<0.01

Table 18: **Overcrowding and Performance of Quantitative Funds**
Independent variable: performance as 12–months rolling CAPM alphas (*Alphas*) and Sharpe Ratio (*SR*). Dependent variables: E_1, E_2 and *Recessions* are dummies identifying the first two expansions and all recessions in my sample respectively; holdings commonality (*Commonality*) defined as the allocation-weighted average of the percentage of other quantitative funds which hold each stock over time (as described is Section 5.1); and a dummy which assumes a value of 1 when the commonality measure is above the median (*DCommonality*).

	12 *Months Rolling*					
	Alphas			*SR*	*Alphas*	
	CAPM	$3-Factor$	$4-Factor$		*CAPM*	*CAPM*
	(1)	(2)	(3)	(4)	(5)	(6)
Constant	-0.106***	-0.105***	-0.0879***	0.251**	0.00906	-0.0530
	(0.000)	(0.000)	(0.000)	(0.019)	(0.891)	(0.194)
E_1	0.798***	0.404***	0.460***	-0.581*	0.758***	0.768***
	(0.000)	(0.000)	(0.000)	(0.087)	(0.000)	(0.000)
E_2	0.121***	-0.00883	0.00869	-0.0613	0.106***	0.107***
	(0.001)	(0.696)	(0.709)	(0.686)	(0.006)	(0.006)
Recessions	0.207***	-0.00384	-0.0809	-0.532**	0.215***	0.209***
	(0.000)	(0.932)	(0.133)	(0.036)	(0.000)	(0.000)
Commonality					-0.773**	
					(0.028)	
DCommonality						-0.0887**
						(0.040)
Observations	53147	53147	53147	52546	43348	43348

p-values in parentheses: * $p<0.10$, ** $p<0.05$, *** $p<0.01$