INVESTOR SENTIMENT AND STOCK RETURN PREDICTABILITY: THE POWER OF IGNORANCE

Catherine D'Hondt, Patrick Roger

Presses universitaires de Grenoble | « Finance »

2017/2 Vol. 38 | pages 7 à 37

ISSN 0752-6180
ISBN 9782706142130

Investor sentiment and stock return predictability: The power of ignorance[1]

Catherine D'Hondt[2], Patrick Roger[3]

ABSTRACT

Sentiment measures, based on the trading activity of retail investors, carry some predictive power of future market returns. In this paper, we use such a sentiment measure on two samples of approximately 25,000 individual investors, who differ in their choices when answering MiFID questionnaires, especially in terms of their appetite for information and professional recommendations. Our data covers 51 months from January 2008 to March 2012. We show that the sentiment of investors who disregard free information and professional advice is the best predictor of future returns on a long-short portfolio based on size. Our findings remain valid when controlling for investor characteristics like spoken language (French or Dutch), portfolio value and financial literacy. Our results bring evidence that sentiment is essentially driven by underdiversification and narrow framing by retail investors. When shared by many investors, sentiment can generate long-lived mispricing, which is, therefore, difficult to arbitrage.

KEYWORDS: Investor sentiment, underdiversification, information seeking, recommendations, retail investors.

1. Introduction

In the 1980s, retail investors were often identified as noise traders because "they trade on noise as if it were information" (Black, 1986). If markets

1. We are grateful to Anthony Bellofatto, Carole Bernard (Editor), Marie-Hélène Broihanne, Egle Karmaziene, Maxime Merli, Alit Polat, Thomas Renault and Tristan Roger for their helpful comments and suggestions. We thank the two anonymous referees; their comments have been very useful to improve the paper. We also thank the participants of the conferences where the paper was presented: 24th MFS annual Conference, Bucarest 2017, Joint conference of the Academy of Entrepreneurial Finance and Academy of Behavioral Finance (Europe) Conference, Hohenheim 2017, 34th International Conference of the French Finance Association, 2017. The authors are grateful to the online brokerage house for providing the data and to the European Savings Institute (Observatoire de l'Epargne Européenne) for its financial support. Patrick Roger gratefully acknowledges the financial support of the Chair in Behavioral Finance at EM Strasbourg Business School. Any errors are the full responsibility of the authors.

2. Louvain Finance (IMMAQ), Louvain School of Management, Catholic University of Louvain – Mailing address: Chaussée de Binche 151, 7000 Mons, BELGIUM – Tel: +32 (0)65 32 33 39 – Email: catherine.dhondt@uclouvain.be

3. LaRGE Research Center, EM Strasbourg Business School, University of Strasbourg – Mailing address: 61 avenue de la Forêt Noire, 67085 Strasbourg Cedex FRANCE – Tel: +33 (0) 3 68 85 21 56 – Email: proger@unistra.fr

were efficient, it would mean that noise traders do not influence prices but since the end of the 90s, an abundant literature on retail investors has developed, notably initiated by Terrance Odean (1998, 1999), to suggest that this is not the case. The main stylized facts in this literature state that 1) retail investors hold underdiversified portfolios, 2) retail investors narrowly frame their decisions, and 3) their trades are correlated.

The first stylized fact, i.e. underdiversification, is highlighted in Lease et al. (1974) and Blume and Friend (1975). More recent studies confirm that retail investors hold largely underdiversified portfolios (Kelly, 1995; Odean, 1999; Kumar, 2007; Goetzmann and Kumar, 2008; Mitton and Vorkink, 2007; Broihanne et al., 2016), containing less than five stocks on average. From a theoretical point of view, investors' desire to hold positively skewed portfolios (Barberis and Huang, 2008; Brunnermeier et al., 2007; Brunnermeier and Parker, 2005) or investors' solvency constraints (Liu, 2014) may justify such an underdiversification.

The second stylized fact, i.e. narrow framing, means that retail investors evaluate stocks in isolation (Barberis et al., 2006). Contrary to the assumptions of both expected utility theory and Markowitz's portfolio choice theory, retail investors do not consider their portfolio as a whole. Their decisions to buy or sell a given stock are motivated by their expectations for this stock, which is optimism/pessimism about the future return on this specific stock. Narrow framing is accentuated when investors' portfolios contain a very low number of different stocks (Kumar and Lim, 2008).

The third stylized fact, i.e. correlated trading, implies that suboptimal diversification choices of retail investors can move stock prices and partly drive future returns, as illustrated in Dorn et al. (2008). Using a sample of 37,000 clients of a German broker, these authors show that trades of retail investors are systematically correlated. They also find that correlated limit orders have some predictive power of subsequent market returns. In addition, Kumar and Lee (2006) show that stocks with high retail concentration comove more together, than they comove with other stocks.

Considering the three stylized facts together, it turns out that correlated trading by underdiversified and narrowly framed investors can generate persistent mispricing. This makes the construction of a sentiment index that can have a significant predictive power of future returns on specific portfolios, especially relevant. In fact, the trading behavior and the portfolio dynamics of retail investors are good signals to measure the (excessive) optimism/

pessimism of market participants. Optimism/pessimism of investors is often translated in terms of investor sentiment, which is defined by Baker and Wurgler (2007) as “a belief about future cash flows and investment risks that is not justified by the facts at hand”.

In the words of Kahneman (2011), sentiment investors think more with their System 1 (fast and automatic) brains than with their System 2 (slow and effortful) brains, when they decide to purchase a stock. Roughly speaking, the brain of a decision maker typically uses two mental systems, named System 1 and System 2. System 1 is automatic, affective and heuristic-based. It is the way of thinking that allows us to tell immediately if someone is angry, after seeing her facial expression. By contrast, System 2 requires effort and is mainly rule-based. We use it when calculating, for example, the product of 13 and 52. When faced with a decision, System 1 makes an immediate assessment based on a first impression and transfers it to System 2. The latter either accepts the System 1’s assessment or modifies it more or less. Research on decision-making shows that System 2 often accepts the suggestion of System 1 or adjusts it only slightly. Hence, System 1 has a strong impact on most of our decisions, including financial decisions. The problem is that System 1 is an associative machine. It is able to construct the story that best incorporates available information but it will not warn you that some information is missing and that you should look for more information. Kahneman (2011) summarizes this situation as follows : “The measure of success of System 1 is the coherence of the story it manages to create. The amount and quality of data are largely irrelevant… System 1 operates as a machine for jumping to conclusions.” In the same vein, Barberis et al. (2014) argue that “first impressions” are important in the decision-making process of retail investors.

When sentiment investors trade in concert and generate a sizeable proportion of the total trading, it becomes costly and risky for rational arbitrageurs to bet against them (Shleifer and Vishny, 1997). An obvious consequence is a potential mispricing. The Internet bubble at the end of the 90s is the typical example of this kind of situation where euphoria contaminates investors’ decisions and prevents rational arbitrageurs to correct price trajectories.[4]

This paper is based on two building blocks related to the above literature. First, that investor sentiment measures often rely more or less explicitly on

4. Moreover, Baker et al. (2012) show that sentiment is contagious across countries.

the behavior of retail investors. In short, a measure of sentiment that scores high (low) is an indicator of excessive optimism (pessimism) among retail investors. Future returns are, therefore, expected to be low (high). Second, sentiment is essentially driven by System 1 thinking. As a result, we expect that a sentiment indicator built on the behavior of investors who do not look for (free) additional information and professional recommendations, is a better predictor of future returns than the same indicator built on the behavior of investors more eager to collect information and to use professional advice.

In order to capture the dynamics of portfolio diversification of retail investors, we use the market sentiment index (MSI henceforth) developed by Roger (2014). The intuition behind this indicator is very simple: when a retail investor, who holds a small number of different stocks in his/her portfolio (for example, 2 or 3), decides to buy a new stock, his/her main motivation is that he/she is optimistic about the future returns on this stock (typical narrow framing). When a lot of underdiversified retail investors increase (decrease) the number of different stocks in their portfolios, they are optimistic (pessimistic) about future returns and sentiment is high(low). As shown by Baker and Wurgler (2007), small caps are more influenced by sentiment than large caps. Roger (2014) shows that the MSI performs better than a number of other sentiment indices in predicting future returns on long-short portfolios based on size. The MSI has also several advantages. First, it can be calculated with any sample of retail investors' portfolios. Second, computing the indicator at a given date t only requires the transition matrix of the process N_t of the number of different stocks in investors' portfolios at date t.[5] Third, the MSI is not contaminated by liquidity considerations.

The second building block of this paper is a proprietary database of 45,085 retail investors' online accounts. First, we have all the investors' trading activity over the period from January 2008 to March 2012. Second, we have the investors' answers to both the Suitability test and the Appropriateness test, which have been required in EU member states since the implementation of MiFID[6] in November 2007. In a nutshell, MiFID requires investment firms to submit questionnaires to their clients in order to determine their

5. One of the most popular sentiment indices is the one built by Baker and Wurgler (2006). Their index is a linear combination of six variables known to be influenced by the optimism/pessimism of investors: the closed-end fund discount, the logarithm of the NYSE share turnover ratio (detrended by the 5-year moving average), the number of IPOs, the average first-day return on IPOs, the share of equity issues in total equity and debt issues and the dividend premium, defined as the log difference in the average market-to-book ratios between dividend payers and non-payers. The sentiment measure is chosen as the first principal component of a Principal Component Analysis of the six variables.

6. MiFID stands for Markets in Financial Instruments Directive.

financial capacity, their financial experience and knowledge, and their investment objectives. Such tests should help firms offer retail investors suitable services and instruments. In particular, assessing *suitability* is required before providing investment advice or portfolio management services while assessing *appropriateness* is required before providing the execution and transmission of orders (what is called 'execution only' in the industry) on complex instruments. By using information available on the MiFID tests, we are able to distinguish A-investors, i.e investors who only filled in the *Appropriateness* test, and S-investors, i.e. investors who filled in both the *Appropriateness* test and the *Suitability* test. Specifically, we consider the Suitability test as a proxy for the investor's appetite for information and professional advice. Since the access to the investment advice tool on the web platform is free (the only cost is the time to complete the questionnaire), A-investors neglect free information and professional advice, compared to S-investors. We conjecture that A-investors are more prone to sentiment trading than S-investors. As a consequence, we expect that the MSI built with the portfolio dynamics of A-investors (S-investors) has a stronger (weaker) predictive power of future returns on a long-short portfolio based on size.

Our results confirm our expectations. The MSI is especially effective when it is based on the subsample of A-investors who self-report a low level of financial literacy. Moreover, the MSI's predictive power becomes even stronger if we isolate the peculiarity of A-investors with respect to S-investors, i.e. using as a predictor of returns the residual of the regression of the MSI of A-investors on the MSI of S-investors. Our findings are robust to a propensity score matching procedure aimed at neutralizing the potential impact of spoken language, wealth and financial literacy.

The paper is organized as follows. Section 2 shortly presents the MSI and its main properties.[7] Section 3 describes our data and provides some descriptive statistics. Section 4 presents the empirical results and Section 5 develops robustness checks. Section 6 concludes.

2. The Market Sentiment Index

As mentioned earlier, the MSI is built on the stylized facts that spring from both underdiversification and narrow framing of retail investors (Roger,

7. See Roger (2014) for the technical details.

2014). The MSI is well-suited to our data because its construction only needs the time series of the number of different stocks held in investors' portfolios. We briefly summarize hereafter the formal definition and the properties of this index. The main mathematical tools of its construction are the properties of Markov chains.

2.1. The Markov chain of diversification levels

Assume that K stocks are traded in the market by a set of I investors over time-periods numbered from 1 to T. N_t is the number of different stocks held by an investor at date t. N_t is a random variable taking values in the set $\{0,\ldots,K\}$.

Let Q_t stands for the one-period transition probability matrix of the stochastic process $(N_t, t = 0,\ldots, T)$. It is defined by:

$$\forall 1 \leq k \leq K, \forall 1 \leq m \leq K, Q_t(k,m) = P(N_{t+1} = m \mid N_t = k) \quad (1)$$

$Q_t(k, m)$ is the probability that an investor's portfolio contains m different stocks at date $t + 1$, knowing that it contained k different stocks at date t. In the empirical part, we assume that $K = 5$ because most of the investors in our sample hold less than 5 stocks (state K receives, then, all portfolios with a number of different stocks greater than or equal to K). The elements of Q_t located above the diagonal are greater than those below the diagonal, when investors have a tendency to increase the number of different stocks in their portfolios. Portfolios become more concentrated when the opposite is true. We should notice that neither the trading volume nor the stock price enter the calculation of the MSI.

2.2. Formal definition of the MSI

If the structure of Q_t signals investor sentiment between dates t and $t + 1$, and is stable over time, the properties of homogeneous Markov chains[8] tell us what happens to the long-term transition matrix which is $Q_\infty = \lim_{n\to\infty} Q_t^n$. In fact, all lines of Q_∞ are equal to the equilibrium distribution of the number of stocks in the portfolios, denoted $N_{\infty,t}$. Roger (2014) defines the MSI as the area below the decumulative distribution function of $N_{\infty,t}$. The intuition is that if investors are optimistic and buy new stocks, the

8. A Markov chain is said homogeneous if Q_t does not depend on t.

structure of Q_t leads to an equilibrium probability distribution of N_∞ that overweights large values of $N_{\infty,t}$. As a consequence, the area above (below) the cumulative (decumulative) distribution of $N_{\infty,t}$ is large (small). MSI_t is formally defined by:

$$MSI_t = \frac{1}{K-1}\sum_{k=1}^{K-1} P(N_{\infty,t} > k) \tag{2}$$

An essential feature of the convergence theorem of Markov chains is that the steady-state equilibrium does not depend on the initial distribution of investors. It turns out that only the changes between $t = 1$ and t are important. Q_t contains useful information about the dynamics of portfolio diversification. During long bullish high-sentiment periods (such as the dotcom bubble), more and more investors enter the market and those already in the market increase their stakes and invest in new stocks, thus increasing diversification.[9] Roughly speaking, $Q_t(k, m) > Q_t(m, k)$ in bullish markets. In bearish markets or recession periods, investors are reluctant to put new money on the table and may sell stocks to finance consumption or because of liquidity needs. Consequently, we expect $Q_t(k, m) \leq Q_t(m, k)$ in bearish markets.[10] The mechanics driving the Markov chain is then clearly linked to the optimism/pessimism of retail investors, assuming that their portfolios are underdiversified and that they narrowly frame their decisions.

3. Data and descriptive statistics

Our study use data from three different sources, Eurofidai[11] and Bloomberg for data on stock prices and a large online Belgian brokerage house for data on individual investors. Our sample covers the period from January 2008 to March 2012 and is based on 45,085 retail investors, who completed 2,333,372 trades across 9,064 different stocks. Two types of information are available in the proprietary database.[12] The first dataset provides detailed information about each trade, i.e. the ISIN code of the instrument, the time-stamp, the trade direction, the executed quantity and

9. Goetzmann and Kumar (2008) note an increase in the mean number of stocks in retail investors' portfolios from 4.28 in 1991 to 6.51 in 1996. In such cases, the elements above the diagonal of the transition matrix increase over time.

10. Some asymmetry, however, may arise due to the disposition effect. As a consequence, reluctance to sell stocks in bearish markets can induce some inertia in Q_t. It turns out that the time-series of the terms on the diagonal of Q may be a relevant measure of pessimism.

11. www.eurofidai.org

12. Each investor is anonymized but registered with a unique code allowing us to select all information relative to any specific investor.

the trade price. We also know the currency in which the trade is executed, which allows us to compute the traded volume in euros.[13] The online brokerage house provides retail investors with access to a large panel of financial instruments. The main traded securities are: stocks, funds, options, warrants, and bonds.[14] The second dataset contains additional information about investors: socio-demographic characteristics such as year of birth, gender and spoken language[15] but also their answers to the MiFID tests.

MiFID came into force in November 2007 across the EU member states.[16] One of its objectives was to increase the level of protection of investment firms' retail clients. Accordingly, MiFID requires investment firms to submit a questionnaire for each client in order to determine his/her financial capacity, his/her financial experience and knowledge, and his/her investment objectives. Such tests should help firms offer their retail clients suitable services and instruments. In particular, assessing *suitability* is required before providing retail investors investment advice or portfolio management services while assessing *appropriateness* is required before providing them with execution and transmission of orders (what is referred to as 'execution only' services) on complex instruments. The way to assess suitability and appropriateness is, however, not constrained and each investment firm is free to devise and organize its own questionnaire(s) provided it abides by some general guidelines.

In our case, the brokerage house made use of two distinct questionnaires for appropriateness and suitability. Assessment of appropriateness mainly requires ensuring that the investor has the necessary experience and knowledge to understand the risks involved in complex financial instruments. In practice, the brokerage house has implemented a specific Appropriateness test (henceforth, A-test) for an exhaustive list of instruments, including shares traded on a non-European market or on a European non-regulated market.[17] In our sample, all retail investors provided answers to this A-test, i.e. 45,085 individuals. Over the sample period, the brokerage

13. When necessary, we use historical exchange rates from the European Central Bank to convert monetary volumes into euros.
14. Only futures cannot be directly traded on the common trading platform. As a result, we do not have data about the trading activity on futures.
15. Belgium has three official languages: French, Dutch and German. French and Dutch are spoken the most. On the online trading platform, investors can choose from three available languages: French, Dutch or English.
16. We refer here to MiFID I (2004/39/EC). MiFID II (2014/65/UE) came into force in January 2018 and then replace the first version of this directive.
17. Such as Multilateral Trading Facilities under the MiFID typology.

house was providing its clients with free access (through the web platform) to an investment advice tool on stocks while it was not offering portfolio management services. To get access to this advice tool, which delivers more detailed information on stocks and professional recommendations, investors had to fill in the Suitability test (henceforth, S-test). In our sample, only 21,738 investors decided to fill in this S-test.

Using information available on the MiFID tests, we are able to distinguish A-investors, i.e investors who only filled in the A-test, and S-investors, i.e. investors who filled in both A-test and S-test. Specifically, we consider the S-test as a proxy for the investor's appetite for information and professional advice. Since the access to the investment advice tool on the web platform is free (the only cost is the time to complete the questionnaire), A-investors neglect free information and professional advice, compared to S-investors.

Over the 51-month period, we count 1,312,519 trades on stocks for S-investors, of which 58% are purchases. For A-investors, we count 1,020,853 trades on stocks, with 57% of them being purchases. In monetary volumes, S-investors (A-investors) trade about €10,065 millions (€9,268 millions), with 52% (51%) of that amount for purchases. For the purpose of our study, we focus on stocks and use information about trading activity to build end-of-month portfolios for each investor. With these data at hand, we compute the monthly average number of stocks held in portfolio as well as the monthly average portfolio value.

Table 1 reports cross-sectional statistics for investors' trading activity. Trade-based measures in Panel A show that S-investors execute more trades than A-investors. This is valid for all the instruments (stocks, options, funds and bonds). S-investors also exhibit a longer trading experience: about 28 months on average, in comparison with 22 months on average for A-investors. In addition, S-investors trade more frequently than A-investors, i.e. the average number of days between two consecutive trades on stocks is smaller for them. In Panel B, the stock portfolio-based variables are consistent with underdiversification. On average, A-investors hold a three-stock portfolio while S-investors hold a six-stock portfolio. This difference is significant at the 1% level and is still valid when we look at the medians that are smaller (1.84 for A-investors and 3 for S-investors). These figures reveal that A-investors hold more underdiversified portfolios than S-investors. When we consider the portfolio value, S-investors hold larger portfolios than A-investors: the monthly average portfolio value is €48,477 for S-investors

Table 1. Descriptive statistics for investors' trading activity

The table reports the cross-sectional mean, median, lower and upper quartiles computed over the sample period respectively for trade-based variables, stock portfolio-based variables, and stock portfolio performance variables. '# trades on stocks' is the number of trades executed on stocks. '# trades on bonds' is the number of trades executed on bonds. '# trades on funds' is the number of trades executed on investment fund shares. '# trades on options' is the number of trades executed on options and warrants. '# trading months' is computed as the number of months between the first trade and the last trade on stocks. 'trade duration' is computed as the median number of days between two trades on stocks. '# stocks in portfolio' is the monthly average number of stocks in portfolio. 'portfolio value' is the monthly average portfolio value in euros. 'value by position' is the monthly portfolio value divided by the monthly number of stocks held in portfolio. 'turnover' is the monthly traded volume divided by the end-of-month portfolio value. 'monthly gross return' is the monthly geometric average return, which captures the profitability of both trades and end-of-month portfolio market values. 'monthly net return' is the monthly geometric average gross return, net of explicit transaction costs. 'A-investors' only filled in the A-test while 'S-investors' filled in both the A-test and the S-test. *, **, *** indicate that means or medians statistically differ at the level of respectively 10%, 5%, or 1%.

	A-investors				S-investors			
Panel A: Trade-based variables	**Mean**	**Q3**	**Median**	**Q1**	**Mean**	**Q3**	**Median**	**Q1**
# trades on stocks	43.73***	34	10***	3	60.38	56	20	7
# trades on bonds	0.08***	0	0	0	0.15	0	0	0
# trades on funds	2.67***	0	0	0	10.10	1	0	0
# trades on options	14.28**	0	0	0	16.57	1	0	0
# trading months	22.03***	38	20***	4	28.02	44	30	12
trade duration (#days)	53.35***	35	8.5***	2	38.37	24.5	7	2
Panel B: Stock portfolio-based variables	**Mean**	**Q3**	**Median**	**Q1**	**Mean**	**Q3**	**Median**	**Q1**
# stocks in portfolio	3.77***	4.33	1.84***	0.69	6.05	7.71	3	1.06
portfolio value (€)	36,956***	17,323	4,523***	997	48,477	28,491	7,539	1,696
value by position (€)	6,898	4,032	1,575***	534	6,084	4,049	1,727	652
turnover	4.11*	0.21	0.07***	0.03	1.81	0.23	0.09	0.04
Panel C: Stock portfolio performance variables	**Mean**	**Q3**	**Median**	**Q1**	**Mean**	**Q3**	**Median**	**Q1**
monthly gross return (%)	–0.47***	0.38	–0.34	–1.41	–0.37	0.38	–0.33	–1.2
monthly net return (%)	–0.65***	0.26	–0.45	–1.59	–0.58	0.24	–0.45	–1.4

and €36,956 for A-investors. For both types of investors, the monthly portfolio values are, however, positively skewed, since the mean value is much larger than the corresponding upper quartile value. This suggests a large dispersion for portfolio value in both subsamples. As for the value by position, we do not observe a statistical difference between the monthly averages, although the medians differ and show a slightly larger value by position for S-investors (€1,727 against €1,575). The monthly average turnover is equal to 4.11 for A-investors and to 1.81 for S-investors. The difference is statistically significant at the 10% level and suggests that A-investors churn, on average, more frequently their portfolio than S-investors. In Panel C, we observe that A-investors earn, on average, a slightly lower monthly return than S-investors but the median returns (that are also negative) do not significantly differ between both subsamples of investors. All in all, S-investors appear to be more sophisticated (or experienced) investors than A-investors. Nevertheless, this apparent higher sophistication (experience) does not lead to a clear advantage in terms of monthly performance, as shown by the comparison of median returns. This observation is consistent with Hoechle et al. (2016), who show that professional advice is not always beneficial to the portfolio performance of retail investors.

Additional descriptive statistics are provided in Table 2. Panel A reports some demographics about investors. The median ages are very close in the two subsamples, i.e. 47 versus 48 years old. 18% of A-investors are female, while we count only 10% of female S-investors. We also observe that Dutch-speaking investors have a small majority in both subsamples (55% or 56%). When we look at the education level, the proportion of investors who report holding a university degree or equivalent is larger for S-investors (73% in comparison with 67%). Information in Panel A of Table 2 reveals statistical differences between the two samples, but these differences are not large enough to explain the results we present in Section 4.

Statistics for the self-reported financial literacy are provided in Panel B of Table 2. For both A-investors and S-investors, we observe a real dispersion across the four levels proposed on the scale. The proportions statistically differ, except for level 1. A larger proportion of A-investors report only a very basic (level 0) knowledge (28% of A-investors versus 22% of S-investors). By contrast, a larger proportion of S-investors select the second level of financial literacy, which states that the investor, “understands the functioning of the financial markets and knows that the fluctuations can be important and that the various sectors and categories of products have different characteristics relating to their revenue, growth and risk profile" (40% of S-investors versus 32% of A-investors). Even if the significant difference on the highest level (level 3) of financial literacy is slight (12% vs. 11%), S-investors tend to self-report a higher financial literacy.

Table 2. Descriptive statistics for investors

Panel A reports demographic characteristics. For each investor, we compute age as the difference between 2012 and the year of birth. '% University degree' refers to the proportion of investors who report they hold a university degree or an equivalent. Panel B displays investors' subjective financial literacy, which refers to the answer to one specific question of the A-test where investors have to self-assess their knowledge of financial markets on a scale of 4 levels. The level 0 is associated with a basic knowledge. The level 1 corresponds to 'a sufficient experience to understand well the importance of a good diversification of risks'. The level 2 states that the investor 'understands the functioning of the financial markets and knows that the fluctuations can be important and that the various sectors and categories of products have different characteristics relating to their revenue, growth and risk profile'. The level 3 refers to an experienced investor 'who manages any aspect of the financial markets'. Panel C exhibits the percentage of trades in the main nationalities of stocks traded. 'A-investors' only filled in the A-test while 'S-investors' filled in both the A-test and the S-test. *, **, *** indicate that proportions or medians statistically differ at the level of respectively 10%, 5%, or 1%.

	A-investors	S-investors
Panel A: Demographics		
Median age	47	48***
% Females	18	10***
% French-speaking	45	44**
% Dutch-speaking	55	56**
% University degree	67	73***
Panel B: Financial literacy		
Level 0	28%	22%***
Level 1	28%	28%
Level 2	32%	40%***
Level 3	12%	11%***
Panel C: Nationality of stocks traded		
Belgian stocks	51%	49%***
US stocks	16%	14%***
French stocks	12%	13%***
Dutch stocks	5%	6%***

Panel C of Table 2 shows the countries where the stocks traded by investors originate. We focus on the main countries that account for more than 80% of trades. No dramatic difference appears between the two subsamples on that aspect, although statistically significant.

To summarize, the differences appearing in Table 2 do not seem large enough to explain the effect we are looking for. It is the reason why the difference in the predictive power of sentiment across subsamples (A-versus S-investors) remains, when we control for some of the variables reported in Table 2.

4. Empirical results

4.1. Correlation analysis

Baker and Wurgler (2007) introduce the "sentiment seesaw" to explain the effect of sentiment on stocks (Figure 1, p133). They show that sentiment can have opposite effects on stock returns, depending on the difficulty of engaging in arbitrage. In high-sentiment periods, stocks that are easy to arbitrage (large stocks) may be undervalued and stocks that are difficult to arbitrage (small stocks) may be overvalued. The opposite appears in low-sentiment periods. As a consequence, we expect small caps to be overvalued in comparison with large caps in high-sentiment periods, the reverse being expected in low-sentiment periods. If this prediction is true, small stocks should have low (high) returns following a high-(low-)sentiment period. A good sentiment measure should help forecast future returns but it should also be more correlated to future returns on small stocks than on large stocks.

We first calculate the sentiment index for the whole sample (*MSI*) and then for the two subsamples. *AMSI* (*SMSI*) denotes the sentiment index calculated with the subsample of A-investors (S-investors). As we are mainly interested in the difference between S-investors and A-investors, we denote *RES* the residual of the regression of *AMSI* on *SMSI*. Our conjecture is that *RES* should be a good sentiment indicator because it extracts the specificity of A-investors who voluntarily base their decisions on less information. These investors are more likely to be either more overconfident or more driven by their System 1 (Barberis et al., 2016). As a consequence, we expect A-investors to be more optimistic (pessimistic) when they are optimistic (pessimistic), compared to S-investors.

Table 3 provides the correlations between the four market sentiment variables (*AMSI*, *SMSI*, *MSI*, *RES*) and economic variables and risk factors, namely the four Fama-French-Carhart factors: market premium (*MKT*), size (*SMB*), book-to-market (*HML*), momentum (*MOM*). We also consider the returns on three size-based portfolios. *Lcaps* (*Mcaps*/*Scaps*) is the return on the value-weighted portfolio built with the tercile of large caps (mid-caps/small caps).[18] The last variable *Small – Big* is the return on the long-short portfolio, which is of particular interest in our study. As usual, this portfolio is long on small caps and short on large caps. Panel A(B) of Table 3 shows contemporaneous (lagged) correlations. For example, the first figure of Panel A is –0.354, which is the empirical correlation between $SMSI_t$, $t = 1,\ldots T$ and MKT_t, $t = 1,\ldots T$ where $t = 1$ corresponds to January 2008 and T to March 2012. The first figure of Panel B is –0.063, which is the correlation between $SMSI_t$, $t = 1,\ldots T - 1$ and MKT_t, $t = 2,\ldots T$.

Panel A of Table 3 shows significant negative contemporaneous correlations between our sentiment measures and portfolio returns. These negative correlations can be interpreted in several ways but are consistent with the disposition effect documented in the literature on retail investors. When prices drop, retail investors tend to keep their losing stocks, or, even worse, to buy new stocks in order to decrease the average buying price. On the up-side, the disposition effect leads people to sell their stocks too early after a price increase. These sales generate a decrease in the sentiment index value. The correlations in columns 1 and 5 to 7 are thus compatible with this usual interpretation. Another interesting observation is that correlations in the last column are positive, even if they are insignificant. This positive sign is compatible with the interpretation that the more investors are optimistic, the more they prefer small stocks, leading to a higher return on small stocks than on large stocks. Because of lack of significance, however, we cannot conclude that there is a contemporaneous relationship between the return on the long-short portfolio and the sentiment index. Consistent with this remark is the absence of a significant correlation between the sentiment measures and the size factor.

Results are different for lagged correlations in Panel B. We observe a difference between large caps and small caps,[19] resulting in significant correlations between *AMSI* or *RES* and the long-short portfolio return.

18. The returns on the three size-based portfolios are directly provided by Eurofidai. These portfolios are the three terciles of the universe of European stocks (16 countries) in the database. We do not restrict the universe to Belgian stocks only because 1) Panel C of Table 2 shows that our retail investors do not focus all their trading activities on Belgian stocks, 2) a subsample of these Belgian retail investors hold on average 58% of foreign stocks (Bellofatto, 2016) and 3) there are almost no large-cap stocks in Belgium. As of December 2017, the BEL20 index included 3 foreign stocks accounting for more than 25% of the index capitalization.

19. There is also a significant difference between mid caps and large caps. To follow the standard methodology, however, we mainly analyze the long-short portfolio of the last column, long on small caps and short on large caps.

Table 3. Correlations between sentiment measures, factors and portfolios over the period from January 2008 to March 2012

Panel A provides contemporaneous correlations and Panel B lagged correlations. The market sentiment indices are *SMSI* (*AMSI*), based on the portfolio diversification dynamics of S-investors (A-investors), and *MSI* calculated with the complete sample of investors. *RES* denotes the residual of the regression of *AMSI* on *SMSI*. The risk factors are the four Fama-French-Carhart factors: the market return *MKT*, the size factor *SMB*, the value factor *HML*, and the momentum factor *MOM*. These four factors come from the Eurofidai database. They are calculated as the corresponding factors on the U.S market. The three portfolios *Lcaps*, *Mcaps* and *Scaps* are also provided by Eurofidai and represent the returns of portfolios based on size terciles (*Lcaps* for large caps, *Mcaps* for midcaps and *Scaps* for small caps). The last column "Small-Big" is the difference between the two portfolio returns *Scaps* – *Lcaps*. *, **, *** indicate statistical significance at the 10%, 5%, or 1% levels respectively.

	MKT	SMB	HML	MOM	Lcaps	Mcaps	Scaps	Small-Big
Panel A: Contemporaneous correlations-January 2008-March 2012								
SMSI	–0.354**	0.096	–0.107	0.051	–0.331**	–0.254*	–0.241	0.150
AMSI	–0.579***	0.188	–0.350**	0.260**	–0.564***	–0.465***	–0.437***	0.200
MSI	–0.475***	0.142	–0.231	0.150	–0.455***	–0.366**	–0.345**	0.176
RES	–0.633***	0.230	–0.526***	0.431***	–0.634***	–0.549***	–0.513***	0.182
Panel B: Lagged correlations-January 2008-March 2012								
SMSI	–0.063	–0.162	0.166	–0.099	–0.030	–0.053	–0.083	–0.107
AMSI	–0.160	–0.291**	–0.000	0.054	–0.128	–0.231*	–0.257**	–0.267**
MSI	–0.108	–0.230*	0.090	–0.033	–0.075	–0.138	–0.166	–0.188
RES	–0.222*	–0.338**	–0.222	0.234	–0.203	–0.370***	–0.379***	–0.367***

On the contrary, the sentiment measure based on S-investors (*SMSI*) is correlated neither with future returns on the long-short portfolio, nor with the Fama-French-Carhart factors. Table 3 reveals, therefore, differences between sentiment built on the portfolio dynamics of well-informed investors and sentiment based on the portfolio dynamics of less-informed investors (i.e who disregard a free opportunity to get more information and professional advice).

From this univariate analysis, we can draw the very preliminary (paradoxical?) conclusion that investors with a greater appetite for information are "noise traders" when it comes to measuring sentiment and forecast returns. In a sense, it is a good point in favor of market efficiency. Yet, the fact that: 1) approximately 50% of our sample belongs to the category of A-investors, who disregard free information, and 2) *AMSI* and *RES* help to forecast returns, leads to the reverse conclusion. Informational efficiency is far from being satisfied. This preliminary analysis shows that a multivariate analysis is useful to conclude whether sentiment can really forecast returns or whether it is only a combination of the usual risk factors.

4.2. Multivariate analysis

Our conjecture is twofold. First, *RES* is the best predictor of future returns among the four sentiment measures (*SMSI*, *AMSI*, *MSI* and *RES*). Second, *SMSI* is a worse predictor than *AMSI*. To address this conjecture, we compare the performance of the four-market sentiment indexes as predictors of future returns on a long-short portfolio based on size.

As mentioned above, the "sentiment seesaw" implies that the long-short portfolio return is high after low-sentiment periods and low after high-sentiment periods (Baker and Wurgler (2007)). When regressing the return of the long-short portfolio on sentiment indexes, we expect a negative sign for the coefficient of the lagged sentiment measures. For that purpose, we replicate the methodology of Baker and Wurgler (2006) that contains two steps, wherein the dependent variable is $R_{Smallcaps,t} - R_{Largecaps,t}$ and $R_{Smallcaps,t}$ ($R_{Largecaps,t}$) is the return on a value-weighted portfolio built with the tercile of small (large) European stocks.

In the first step, we estimate the following regression equation:

$$R_{Smallcaps,t} - R_{Largecaps,t} = \alpha + \beta_s . \, Sentiment_{t-1} + \varepsilon_t \tag{3}$$

where $Sentiment_t$ is the sentiment index for month t and may be either *MSI*, *SMSI*, *AMSI* or *RES*.

In the second step, we control for Fama-French-Carhart factors with the following regression model:

$$R_{Smallcaps,t} - R_{Largecaps,t} = c + \beta_s.\ Sentiment_{t-1} + \beta_X X_t + \varepsilon_t \tag{4}$$

The vector X of control variables includes the market factor (*MKT*) and the two Fama-French-Carhart factors, namely the book-to-market factor (*HML*) and the momentum factor(*MOM*). The size factor is not included in the equation because it is almost perfectly correlated with the dependent variable. The data for these factors come from the Eurofidai database.

Table 4. Coefficients of sentiment when regressing the returns of a long-short portfolio based on size, on sentiment measures

Panel A gives the coefficient of sentiment in the simple regression: $R_{Smallcaps,t} - R_{Largecaps,t} = \alpha + \beta_s.\ Sentiment_{t-1} + \varepsilon_t$. Panel B provides the same coefficient when controlling for Fama-French factors and the Carhart momentum factor. The regression equation is then : $R_{Smallcaps,t} - R_{Largecaps,t} = \alpha + \beta_s.\ Sentiment_{t-1} + \beta_X X_t + \varepsilon_t$ where the matrix X includes the market factor and the two Fama-French-Carhart factors *HML*, *MOM* (*SMB* is not included because the long-short portfolio is based on this criterion). The sentiment indexes are the four measures *SMSI*, *AMSI*, *MSI* and *RES*. *AMSI* (*SMSI*) is calculated with the sample of A-investors (S-investors) who filled in the appropriateness test (the two tests, appropriateness and suitability). *MSI* is calculated with the complete sample. *RES* is the residual of the regression of *AMSI* on *SMSI*. When sentiment is not considered in the controlled equation, the adjusted R^2 of the regression is 0.073. *, **, *** indicate statistical significance at the 10%, 5%, or 1% levels, respectively.

	SMSI	AMSI	MSI	RES
	Panel A: Equation (3) without controls			
β_s	–0.060	–0.118**	–0.095	–0.310***
t-stat	–0.703	–2.320	–1.384	–3.241
p-val	0.485	0.024	0.172	0.002
$\bar{R}^2$	–0.009	0.052	0.015	0.117
	Panel B: Equation (4) with controls			
β_s	–0.088	–0.137***	–0.120*	–0.336***
t-stat	–1.211	–2.807	–1.967	–3.398
p-val	0.232	0.007	0.055	0.001
$\bar{R}^2$	0.077	0.152	0.110	0.213

Panel A (B) of Table 4 provides the regression coefficients of Equation 3 (4) without (with) control for the Fama-French-Carhart factors. The expected negative sign for the sentiment coefficient (β_s) appears in the two models. Consistent with the extant literature, periods of high (low) sentiment are followed by low (high) returns on the long-short portfolio, even after controlling for the market, book-to-market and momentum factors. In addition, the most significant coefficients are those of *AMSI* and *RES* while the *SMSI* coefficient is never significant. These findings indicate that sentiment measures built on the portfolio dynamics of A-investors are much better predictors of future returns, compared to measures based on the well-informed S-investors (and even in the case where they are more active in terms of trading volume). This means that many trades executed by *S*-investors are not informative about their portfolio dynamics when characterized by the MSI. It is not really surprising because trades completed by these investors who have access to information and professional recommendations are more likely to be motivated by portfolio management concerns. Tables 1 and 2 show that *S*-investors hold more diversified portfolios, are more financially literate on average, and more wealthy than *A*-investors. This could suggest that a proportion of their trades does not change the number of different stocks they hold because they are motivated by portfolio adjustments. Such trades do not move the sentiment index.[20]

Finally, Table 4 shows that the adjusted R^2 is largely improved by the introduction of sentiment in the controlled version of the model, especially when we consider the variable *RES*. In that case, the adjusted R^2 reaches 21%.

5. Robustness tests

In order to check the robustness of our previous results, we perform three types of tests. First, we take into account the possible autocorrelation of sentiment measures that could overestimate the predictive power of sentiment. Second, as we interpreted the difference between A-investors and S-investors observed in Table 4 in terms of appetite for information and differences in decision processes, we have to rule out some alternative explanations. In particular, we test whether the differences for sentiment between A-investors and S-investors are (partially) driven by differences respectively in cultural background, in financial literacy, or in wealth. Finally, we build matched samples of investors on the three aforementioned

20. We do not say that sentiment is absent for these trades, but only that another sentiment index like a buy-sell imbalance measure should be used to take into account the sentiment of *S*-investors.

aspects and replicate the main analysis to check whether *RES* remains a good predictor of future returns.

5.1. Autocorrelation of sentiment measures

As mentioned in Roger (2014), the regressions in Table 4 can produce biased estimators of β_s when *Sentiment* is an autoregressive process. The predictive power of *Sentiment* could, therefore, be overstated. We use the method of Stambaugh (1999) and Amihud and Hurvich (2004) to reduce the bias of the estimator.

Let the regression of $Sentiment_t$ on $Sentiment_{t-1}$ be written as:

$$Sentiment_t = \theta + \rho Sentiment_{t-1} + \nu_t \tag{5}$$

1) The estimate $\hat{\rho}$ is corrected as follows

$$\hat{\rho}^c = \hat{\rho} + \frac{1+3\hat{\rho}}{n} + \frac{3(1+\hat{\rho})}{n^2} \tag{6}$$

where $n = 51$ when the regression covers the entire sample period.

2) The residuals of regression are estimated using $\hat{\rho}^c$ and denoted by $\nu^c = (\nu_t^c, t = 1,..n)$. The vector ν^c is introduced in regression (3), which becomes

$$R_{Smallcaps,t} - R_{Largecaps,t} = \alpha + \phi\nu_t^c + \beta_s.\ Sentiment_{t-1} + \varepsilon_t \tag{7}$$

In the controlled case, the equation writes:

$$R_{Smallcaps,t} - R_{Largecaps,t} = c + \beta_s.\ Sentiment_{t-1} + \beta_X X_t + \phi\nu_t^c + \varepsilon_t \tag{8}$$

Finally, the corrected standard error of β_s is

$$\widehat{SE}^c(\beta_s) = \sqrt{\hat{\phi}^2 Var(\hat{\rho}^c) + \widehat{SE}^2(\beta_s)} \tag{9}$$

where $\widehat{SE}^c(\beta_s)$ is used to calculate the significance of the estimator of β_s in Table 5.

A comparison of Tables 4 and 5 leads to two main comments. First, the significant coefficients are the same in both tables and the significance levels are comparable. In particular, the coefficients of *RES* are virtually unchanged. This is not surprising because *RES* is already the residual of the regression of *AMSI* on *SMSI*. As a consequence, the autocorrelation of this sentiment index is much lower than that of the other sentiment indicators. The level of autocorrelation varies between 0.281 for *RES* to 0.507 for *SMSI*. The main difference between Tables 4 and 5 lies in the adjusted R^2. In most cases, the adjusted R^2 is higher in Table 5 because one more explanatory variable (v_t) appears in the model. The only exception concerns the sentiment indicator *RES* in the controlled case, for which the adjusted R^2 decreases. This observation is consistent with our previous remark concerning the low autocorrelation of this variable.

Table 5. Coefficients of sentiment when regressing the returns of a long-short portfolio based on size, on sentiment measures, using the reducing-bias technique of Amihud and Hurvich (2004)

Panel A gives the coefficient of sentiment in the simple regression: $R_{Smallcaps,t} - R_{Largecaps,t} = \alpha + \beta_s \cdot Sentiment_{t-1} + \phi \nu_t + \varepsilon_t$. Panel B provides the same coefficient when controlling for Fama-French factors and the Carhart momentum factor. The regression equation is then: $R_{Smallcaps,t} - R_{Largecaps,t} = c + \beta_s \cdot Sentiment_{t-1} + \beta_X X_t + \phi \nu_t + \varepsilon_t$ where the matrix X includes the market factor and the two Fama-French-Carhart factors *HML*, *MOM* (*SMB* is not included because the long-short portfolio is based on this criterion). The sentiment indexes are the four measures *SMSI*, *AMSI*, *MSI* and *RES*. *AMSI* (*SMSI*) is calculated with the sample of A-investors (S-investors) who filled in the appropriateness test (the two tests, appropriateness and suitability). *MSI* is calculated with the complete sample. *RES* is the residual of the regression of *AMSI* on *SMSI*. When sentiment is not considered in the controlled equation, the adjusted R^2 of the regression is 0.073. *, **, *** indicate statistical significance at the 10%, 5%, or 1% levels, respectively.

	SMSI	AMSI	MSI	RES
Panel A: Equation (7) without controls				
β_s	–0.049	–0.113*	–0.086	–0.307***
t-stat	–0.541	–1.977	–1.141	–3.030
p-val	0.590	0.054	0.259	0.004
$\bar{R}^2$	0.040	0.122	0.088	0.134
Panel B: Equation (8) with controls				
β_s	–0.077	–0.125**	–0.106	–0.329***
t-stat	–0.998	–2.460	–1.642	–3.173
p-val	0.324	0.018	0.108	0.003
$\bar{R}^2$	0.086	0.165	0.130	0.196

5.2. Language and cultural background

The two main languages spoken in Belgium are French and Dutch. Our sample is well balanced, as shown in Table 6. 20,036 investors are French-speakers and 25,049 are Dutch speakers. In the Belgian population, there are twice as many people in the Flemish region (populated by Dutch speakers) than in the Walloon region (populated by French speakers).[21] There are a number of cultural and economic differences between these two communities. Such differences could imply a different appetite for information and/or a different attitude with respect to professional advice, etc. For example, a report by the Council of Europe in 2001 states (Council of Europe, 2001): "Taking the form of a triangle tilted from north-west to south-east, Belgium is traversed, on an east-west line running almost through its center, by one of Europes oldest "cultural frontiers". This corresponds roughly to the line at which Julius Caesars armies stopped in their conquest of Gaul in the 1st century BC. Latin exercised a decisive influence south of this line but remained secondary, north of it."

Table 6. Subsamples of investors for the robustness checks
We report in the table the number of investors for each subsample under scrutiny. For financial literacy, we use the specific question of the A-test where investors have to self-assess their knowledge of financial markets on a scale of 4 levels. In the 'Low financial literacy' subsample, we consider the investors who choose the first two levels while the investors who select the other two levels are put in the 'High financial literacy' subsample. To discriminate investors on their portfolio market value, we first compute the cross-sectional median monthly portfolio value across all investors. The investors whose monthly average portfolio value is smaller (larger) than the cross-sectional median monthly portfolio value are put in the 'Small portfolio value' ('Large portfolio value') subsample. A-investors only filled in the A-test while S-investors filled in both the A-test and the S-test.

	A-investors	S-investors
French-speaking	10,489	9,547
Dutch-speaking	12,858	12,191
Low financial literacy	13,131	10,836
High financial literacy	10,216	10,902
Small portfolio value	12,633	9,910
Large portfolio value	10,714	11,828

21. See 'Regional Accounts' on the website of the National Bank of Belgium, http://stat.nbb.be.

In this subsection, we test whether the difference between *AMSI* and *SMSI* could be driven by a disequilibrium between French and Dutch speakers. Our test consists of dividing our sample into four sub-categories denoted *A_FR*, *S_FR*, *A_NL*, *S_NL* respectively. The suffix FR stands for French and NL for Dutch.[22] The prefixes *A* and *S* retain the same meaning as before.

We recalculate the sentiment indexes for the four subsamples and the associated variables *RES_FR* and *RES_NL*. Results are provided in the top six rows of Table 7. The left (right) part of the table provides the uncontrolled (controlled) regression results. For the two subsamples (French speakers and Dutch speakers), the variables *RES* are highly significant in both versions of the model (uncontrolled and controlled). Despite the different cultural and economic background, we observe a large difference between A-indexes and S-indexes. The S-index is never significant, while the A and RES coefficients are always significant. In particular, the *t*–statistics of the *RES* coefficients vary between –2.197 and –4.067, showing that this variable is highly significant in all subsamples. Nevertheless, we observe a difference between French speakers and Dutch speakers. For French speakers, the coefficients of *A_FR* are more significant than the *RES_FR* coefficients. The adjusted R^2 are also higher with the *A_FR* variable than with the *RES_FR* variable. For example, in the controlled case, the adjusted R^2 with *A_FR* is equal to 17.1% but it is only 13.1% with the *RES_FR* variable.[23] With the subsample of Dutch speakers, the significance of *RES_NL* is much stronger than that of *A_NL*. In the controlled case, the adjusted R^2 with *A_NL* is equal to 13.5% but it goes up to 20.8% with the *RES_NL* variable. Beyond these minor differences, it appears that sentiment is more present in the portfolio dynamics of A-investors, as expected.

22. The Flemish language is close to the Dutch language and the Flemish part of Belgium is essentially in the north of the country, closer to Netherlands than to France. See for example http://ies.berkeley.edu/enews/articles/flemishlanguage.html.

23. In both cases, however, the adjusted R^2 is well larger than when sentiment is not included.

Table 7. Regression coefficients of sentiment measures on subsamples

S, *A* denote sentiment measures built with, respectively *S*-investors and *A*-investors. *RES* denotes the residual of the regression of the *A*-based measure on the *S*-based measure. The suffixes *FR* and *NL* denote the main language spoken by the investors (*FR* for French or *NL* for Dutch). For example, *A_FR* is the sentiment measure built on the portfolio dynamics of the subset of *A*-investors who are French speakers. The suffixes *LFL* and *HFL* identify the self-reported financial literacy (low for *LFL*, high for *HFL*). *LPV* and *SPV* identify the subsamples based on portfolio value. *LPV* (*SPV*) means large (small) portfolio value. *, **, *** indicate statistical significance at the 10%, 5%, or 1% levels respectively.

	NO CONTROL				CONTROL			
Variable	Coefft	t-stat	p-value	R^2	Coefft	t-stat	p-value	R^2
S_FR	–0.091	–1.234	0.224	0.065	–0.106	–1.682	0.100	0.109
A_FR	–0.108**	–2.285	0.027	0.136	–0.116**	–2.664	0.011	0.171
RES_FR	–0.184**	–2.197	0.033	0.082	–0.188**	–2.215	0.032	0.131
S_NL	–0.006	–0.066	0.948	–0.010	–0.035	–0.464	0.645	0.047
A_NL	–0.101*	–1.743	0.088	0.077	–0.117**	–2.349	0.023	0.135
RES_NL	–0.274***	–3.164	0.003	0.111	–0.327***	–4.067	0.000	0.208
S_LFL	–0.029	–0.373	0.711	0.037	–0.055	–0.864	0.392	0.089
A_LFL	–0.106**	–2.312	0.025	0.171	–0.116***	–2.828	0.007	0.205
RES_LFL	–0.249***	–3.050	0.004	0.164	–0.268***	–3.697	0.001	0.235
S_HFL	–0.073	–0.771	0.444	0.001	–0.092	–1.078	0.287	0.056
A_HFL	–0.087	–1.189	0.240	0.020	–0.102	–1.631	0.110	0.085
RES_HFL	–0.122	–1.062	0.294	–0.012	–0.129	–1.300	0.200	0.066
S_LPV	–0.024	–0.166	0.868	0.019	–0.033	–0.267	0.790	0.063
A_LPV	–0.119	–1.638	0.108	0.058	–0.109*	–1.698	0.097	0.095
RES_LPV	–0.262**	–2.606	0.012	0.055	–0.242**	–2.665	0.011	0.107
S_SPV	–0.076	–0.937	0.354	0.044	–0.137*	–1.997	0.052	0.123
A_SPV	–0.123*	–1.890	0.065	0.136	–0.210***	–3.629	0.001	0.247
RES_SPV	–0.262	–1.442	0.156	0.040	–0.369*	–1.935	0.059	0.163

5.3. Financial literacy

Panel C of Table 2 shows that we have more A-investors than S-investors in the two levels of low financial literacy (henceforth, *LFL*), the reverse being true in the other higher levels of financial literacy (henceforth, *HFL*). To keep only two subsamples with respect to literacy, we aggregate in Table 6 levels 0 and 1 to define the *LFL* subsample, and 2 and 3 to define the *HFL* subsample.

We replicate the analysis of subsection 5.2 for the four variables *S_LFL*, *A_LFL*, *S_HFL*, *A_HFL* and the two variables *RES_LFL* and *RES_HFL*.

Results are reported in the middle of Table 7. For the subsample of low literate investors, we get the same result as before. The sentiment measure *RES_LFL* is a very good predictor of future returns with *t*-statistics respectively equal to –3.05 (–3.697) in the uncontrolled (controlled) case. The coefficients of the measure based on A-investors are also significant while the coefficients for S-investors are not significant with *t*-statistics lower than 1 in absolute value. The findings are different for investors who report high financial literacy. No significant difference appears between the two subsamples of A- and S-investors. One explanation of this result could be related to Panel C of Table 2. More than 10% of investors (either A- or S-investors) self-report the highest level of financial literacy, which corresponds to investors "who manage any aspect of financial markets". Most academic readers specialized in finance would not choose this level because they know that it is impossible to reach such a level of competence in globalized and complicated markets. Hence, choosing not to fill in the S-test for these investors could also mean that they are already very well informed and benefit from professional advice elsewhere. It is, then, not so surprising that no significant difference emerges between the highly literate A-investors and S-investors.

5.4. Portfolio value

The portfolio dynamics, and, in particular, its diversification degree, are influenced by the portfolio value in a mechanical way. Ceteris paribus, the number of stocks in a portfolio is lower for an investor who has only a few hundred or thousand euros to invest, compared to an investor whose portfolio is worth one or two hundred thousands euros. Table 1 shows that the portfolio value of 25% of A-investors (S-investors) is worth less than €997 (€1,696). This could suggest that the difference between the sentiment index of A-investors and S-investors can be significant for large portfolio values (above the median for example) but not for small portfolio values (under the median). Portfolios worth a few hundred euros are too constrained to generate a significant difference between A-investors and S-investors. In particular, for the most constrained investors, the purchase of a stock is often financed by the sale of another stock, keeping unchanged the number of different stocks in portfolio.

We again replicate the analysis of subsection 5.2 in order to address the impact of portfolio value. Results are presented in the bottom rows in

Table 7. For the subsample of large portfolios (*LPV*), only the coefficients of *RES_LPV* are still significant in both models. The coefficients of *A_LPV* are only significant in the model with control. As expected, the results for the subsamples of low portfolio values (*SPV*) are mixed. In particular, all the coefficients are significant in the model with control but significance is the highest for the variable *A_SPV*.

5.5. A-investors matched with S-investors

Finally, we use propensity score matching (Rosenbaum and Rubin, 1983) to simultaneously control for language, financial literacy and portfolio value. Our purpose is to select two groups of "comparable" A-investors and S-investors, i.e. investors who mainly differ by their appetite for information and professional recommendations.

For this purpose, we compute the propensity scores using a logit model wherein the dependent variable, Y_i, is a binary variable that equals 1 if the investor *i* filled in the S-test and 0 otherwise. The probability of being a S-investor is conditioned on a set of regressors, *X*, and is given by $Prob[Y = 1 \mid X] = \wedge(x / b)$, where $\wedge(.)$ is the logistic cumulative distribution function. The set of regressors is made of an intercept, one dummy for the language, $N - 1$ dummies (that is 3) for the level of self-reported financial literacy, and the log value of one plus the investor's monthly average portfolio value. Our logit model is then the following:

$$Y_i = \alpha + \beta_1 NL_i + \beta_2 FL1_i + \beta_3 FL2_i + \beta_4 FL3_i + \beta_5 LOG(1 + MPV_i) \quad (10)$$

where NL_i is equal to one when the investor *i* is Dutch-speaking, $FL1_i$ is set to one when the investor *i* selects the level 1 of financial literacy, $FL2_i$ is set to one when the investor *i* selects the level 2 of financial literacy, $FL3_i$ is set to one when the investor *i* selects the level 3 of financial literacy and MPV_i refers to the monthly average portfolio value computed for investor *i*.

Table 8 reports the results of the logit model. All the parameter estimates are statistically significant, except the dummy variable, for the highest level of financial literacy. When looking at the odds ratios, we observe a positive relationship between the probability of being a S-investor and all the regressors (except the intercept). Investors who self-report a higher literacy (level 1 or 2) or investors who hold larger portfolios, are more likely to display a higher appetite for financial information. This relationship is also present for Dutch-speaking investors, even if it is somewhat weaker.

Based on the propensity scores estimated by our logit model, we then match S-investors with A-investors using the caliper matching method with replacement. This approach is similar to the nearest available neighbor matching method, with an additional restriction to avoid bad matches. Each treated unit (S-investor in our case) is selected to find its closest control match (A-investor in our case) based on the propensity scores but the control's propensity score is required to be within a certain radius (named caliper).[24] This restriction implies that it is possible that some S-investors cannot be matched to a real "comparable" A-investor. We end up with a subsample of 7,929 S-investors and a corresponding sub-sample of 7,929 matched A-investors.[25]

Table 8. Results of the logit model

This table reports the results for the logit model wherein the dependent variable, Y_i, is a binary variable that equals 1 if the investor *i* filled in the S-test and 0 otherwise. NL_i is equal to one when the investor *i* is Dutch-speaking, $FL1_i$ is set to one when the investor *i* select the level 1 of financial literacy, $FL2_i$ is set to one when the investor *i* select the level 2 of financial literacy, $FL3_i$ is set to one when the investor *i* select the level 3 of financial literacy, and MPV_i refers to the monthly average portfolio value computed for the investor *i*. The odds ratio for a given explanatory variable is the exponential of its estimated coefficient. When the independent variable is continuous, the odds ratio measures how the probability of success changes if the variable increases by one unit (from x to x+1). For a binary variable, the odds ratio assesses how the probability that the event will occur changes when the variable goes from zero to one. If the odds are greater (lower) than one, then the event is more (less) likely to happen. *, **, *** indicate statistical significance at the 10%, 5%, or 1% levels, respectively.

Independent variables	Parameter estimates	Odds Ratios	95% Wald	Confidence Limits
Intercept	–1.0533***			
FL1	0.1909***	1.210	1.149	1.275
FL2	0.3710***	1.449	1.379	1.523
FL3	0.0121	1.012	0.945	1.084
NL	0.0366*	1.037	0.999	1.077
LOG(1+MPV)	0.0908***	1.095	1.085	1.105

Table 9 provides a comparison of both subsamples for the control variables under scrutiny. Matched A-investors and S-investors no longer differ on portfolio value, spoken language and financial literacy.[26]

24. We set the caliper at 10^{-5}.
25. Since we allow replacement, we have, in fact, 6,538 matched A-investors, some of them being matched to several S-investors.
26. In addition to this univariate approach, we run the logit model on our matched subsamples and the results are no longer significant, except for the highest level of financial literacy at 10% level. The detailed results are available upon request.

We recalculate the sentiment indexes on the matched subsamples and duplicate the methodology used before. Table 10 reports the results. Panels A and B (C and D) refer to the regressions without (with) adjustment for autocorrelation. As the results are very close, we focus on Panels C and D. Moreover, we make comparisons with Table 5, which is devoted to the initial subsamples. First, the coefficients of *AMSI* and *RES* remain always significant at the 5% level on matched subsamples, even if the matched A-investors are comparable to the S-investors in terms of wealth, financial literacy and language. In particular, the matched sample of A-investors is more financially literate than the general A-sample, as we can see when comparing Table 2 and Table 9. The significance of the coefficient of *AMSI* is almost unchanged in the matched subsample and the coefficient of *RES* decreases in absolute value but is still strongly significant with *t –stats* around 3 in absolute value. Concerning the adjusted R^2, we observe an increase in the matched subsamples for *AMSI* and a slight decrease for *RES*.

Table 9. Mean comparisons for control variables

This table reports the results of mean comparisons for control variables between S-investors and matched A-investors. Investors' matching is based on the results of the logit model presented in Table 8, using a propensity score matching (with replacement) to simultaneously control for language, financial literacy, and portfolio value. NL_i is equal to one when the investor *i* is Dutch-speaking, $FL0_i$ is set to one when the investor *i* select the level 0 of financial literacy, $FL1_i$ is set to one when the investor *i* select the level 1 of financial literacy, $FL2_i$ is set to one when the investor *i* select the level 2 of financial literacy, $FL3_i$ is set to one when the investor *i* select the level 3 of financial literacy, and MPV_i refers to the monthly average portfolio value computed for the investor *i*. *, **, *** indicate that mean differences are statistically significant at the level of respectively 10%, 5%, or 1%. 'A-investors' only filled in the A-test while 'S-investors' filled in both the A-test and the S-test.

	Matched A-investors	S-investors
NL	0.5564	0.5614
FL0	0.2143	0.2073
FL1	0.3179	0.3293
FL2	0.3592*	0.3462
FL3	0.1086*	0.1172
MPV	28,879	27,287
LOG(1+MPV)	8.6847	8.7107

Table 10. Coefficients of sentiment when regressing the returns of a long-short portfolio based on size, on sentiment measures calculated over the period February 2008 to March 2012 on matched samples

Panels A and B provide the unadjusted results calculated as in Table 4. Panels C and D provide results adjusted for autocorrelation of sentiment measures, as in Table 5. Panel A (C) gives the coefficient of sentiment in the simple regression: $R_{Smallcaps,t} - R_{Largecaps,t} = \alpha + \beta_s . Sentiment_{t-1} + (+\phi\nu_t) + \varepsilon_t$. Panel B (D) provides the same coefficient when controlling for Fama-French factors and the Carhart momentum factor. The regression equation is then: $R_{Smallcaps,t} - R_{Largecaps,t} = \alpha + \beta_s . Sentiment_{t-1} + \beta_X X_t + (+\phi\nu_t) + \varepsilon_t$ where the matrix X includes the market factor and the three Fama-French-Carhart factors (size is not included because the portfolio is based on this criterion). The sentiment indexes are the measures *SMSI*, *AMSI*, and *RES*. *SMSI* (*AMSI*) is calculated with the sample of S-investors who filled in the two tests (the appropriateness test). *RES* is the residual of the regression of *AMSI* on *SMSI*. When sentiment is not considered in the controlled equation, the adjusted R^2 of the regression is 0.073. *, **, *** indicate statistical significance at the 10%, 5%, or 1% levels, respectively.

	SMSI	AMSI	RES
	Panel A: Matched sample-Equation (3) without controls		
β_s	–0.007	–0.119**	–0.222***
t-stat	–0.083	–2.239	–2.879
p-val	0.933	0.030	0.005
$\bar{R}^2$	–0.021	0.065	0.134
	Panel B: Matched sample-Equation (4) with controls		
β_s	–0.058	–0.124**	–0.207***
t-stat	–0.842	–2.563	–3.038
p-val	0.404	0.014	0.003
$\bar{R}^2$	0.062	0.147	0.180
	Panel C: Matched sample-Equation (7) without controls		
β_s	0.002	–0.118**	–0.222***
t-stat	0.019	–2.179	–2.835
p-val	0.984	0.034	0.006
$\bar{R}^2$	–0.007	0.084	0.119
	Panel D: Matched sample-Equation (8) with controls		
β_s	–0.044	–0.121**	–0.212***
t-stat	–0.593	–2.509	–3.196
p-val	0.556	0.016	0.003
$\bar{R}^2$	0.053	0.132	0.163

6. Conclusion

Thanks to a proprietary database, we have the opportunity to distinguish two categories of retail investors on the basis of their appetite for information and professional recommendations. A-investors, who filled in only an appropriateness test, neglect free information and professional recommendations. On the contrary, S-investors, who also filled in a suitability test, get free access (through the web platform) to an investment advice tool that delivers more detailed information on stocks and professional recommendations. We produce evidence that the MSI built on the portfolio dynamics of A-investors is a better predictor of returns on a long-short portfolio based on size than the corresponding MSI based on the portfolio dynamics of S-investors. In addition, the sentiment indicator, that takes into account the peculiarities of A-investors, i.e. the residual of the regression of the MSI of A-investors on the MSI of S-investors, does perform the best in terms of forecasting future returns.

Our findings are robust to several variations and controls. In particular, they remain valid on subsamples of Dutch-speaking or French-speaking investors, despite the cultural and economic differences between of these two communities in Belgium. Our results are further strengthened when we focus on low-literate investors. They are also valid for the subsample of investors in the upper half of wealth. Using a propensity score matching procedure to simultaneously control for the potential impact of the aforementioned variables on our findings, we show that the portfolio dynamics of A-investors still delivers a better MSI to forecast future returns on a long-short portfolio based on size.

All these results produce empirical evidence that market sentiment is related to correlated trading by underdiversified and narrowly framed investors. When shared by many retail investors, market sentiment can generate long-lived mispricing that is, therefore, difficult to arbitrage.

References

Y. Amihud and C. M. Hurvich. Predictive regressions: A reduced-bias estimation method. *Journal of Financial and Quantitative Analysis*, 39(4):813-841, 2004.

M. Baker and J. Wurgler. Investor sentiment and the cross-section of stock returns. *Journal of Finance*, 61(4):1645-1680, 2006.

M. Baker and J. Wurgler. Investor sentiment in the stock market. *Journal of Economic Perspectives*, 21(2):129-151, 2007.

M. Baker, X. Pan, and J. Wurgler. The effect of reference point prices on mergers and acquisitions. *Journal of Financial Economics*, 106:49-71, 2012.

N. Barberis and M. Huang. Stocks as lotteries: The implications of probability weighting for security prices. *American Economic Review*, 98(5):2066-2100, 2008.

N. Barberis, M. Huang, and R. H. Thaler. Individual preferences, monetary gambles, and stock market participation: A case for narrow framing. *American Economic Review*, 96(4):1069-1090, 2006.

N. Barberis, A. Mukherjee, and B. Wang. First impressions: System 1 thinking and stock returns. Conference paper, AFA 2014, Philadelphia, 2014.

N. Barberis, A. Mukherjee, and B. Wang. Prospect theory and stock returns: An empirical test. *Review of Financial Studies*, 29(11):3068-3107, 2016.

A. Bellofatto. How does language impact foreign investing in a multilingual country? Working paper, available at ssrn: https://ssrn.com/abstract=3001825, Universit catholique de Louvain - Louvain School of Management (IMMAQ), 2016.

F. Black. Noise. *Journal of Finance*, 41(3):529-543, 1986.

M. E. Blume and I. Friend. The asset structure of individual portfolios and some implications for utility functions. Rodney L. White Center for Financial Research Working Papers 10-74, Wharton School Rodney L. White Center for Financial Research, 1975.

M.-H. Broihanne, M. Merli, and P. Roger. Diversification, gambling and market forces. *Review of Quantitative Finance and Accounting*, 47(1):129-157, 2016.

M. K. Brunnermeier and J. A. Parker. Optimal expectations. *American Economic Review*, 95(4):1092-1118, 2005.

M. K. Brunnermeier, C. Gollier, and J. A. Parker. Optimal beliefs, asset prices, and the preference for skewed returns. *American Economic Review*, 97(2):159-165, 2007.

Council of Europe. Transversal study: Cultural policy and cultural diversity. Report cc-cult (2001) 3, Council of Europe, 2001.

D. Dorn, G. Huberman, and P. Sengmueller. Correlated trading and returns. *Journal of Finance*, 63(2):885-920, 2008.

W. N. Goetzmann and A. Kumar. Equity portfolio diversification. *Review of Finance*, 12(3):433-463, 2008.

D. Hoechle, S. Ruenzi, N. Schaub, and M. Schmid. The impact of financial advice on trade performance and behavioral biases. *Review of Finance*, forthcoming, 2016.

D. Kahneman. *Thinking, fast and slow*. New York: Farrar, Straus and Giroux, 2011. ISBN 9780374275631.

M. Kelly. All their eggs in one basket: Portfolio diversification of us households. *Journal of Economic Behavior & Organization*, 27(1):87-96, 1995.

A. Kumar. Do the diversification choices of individual investors influence stock returns? *Journal of Financial Markets*, 10(4):362-390, 2007.

A. Kumar and C. M. C. Lee. Retail investor sentiment and return comovements. *Journal of Finance*, 61:2451-2486, 2006.

A. Kumar and S. S. Lim. How do decision frames influence the stock investment choices of individual investors. *Management Science*, 54(6):1052-1064, 2008.

R. C. Lease, W. G. Lewellen, and G. G. Schlarbaum. The individual investor: Attributes and attitudes. *Journal of Finance*, 29(2):413-33, 1974.

H. Liu. Solvency constraint, underdiversification, and idiosyncratic risks. *Journal of Financial and Quantitative Analysis*, 49(2):1-39, 2014.

T. Mitton and K. Vorkink. Equilibrium underdiversification and the preference for skewness. *Review of Financial Studies*, 20(4):1255-1288, 2007.

T. Odean. Are investors reluctant to realize their losses? *Journal of Finance*, 53:1775-1798, 1998.

T. Odean. Do investors trade too much? *American Economic Review*, 89(5):1279-1298, 1999.

P. Roger. The 99% market sentiment index. *Finance*, 35(3):53-96, 2014.

P. R. Rosenbaum and D. B. Rubin. The central role of the propensity score in observational studies for causal effects. *Biometrika*, 70(1):41-55, 1983.

A. Shleifer and R. W. Vishny. The limits of arbitrage. *Journal of Finance*, 52:35-55, 1997.