OBSERVATOIRE DE L'EPARGNE EUROPÉENNE

MIFID questionnaires, financial advice and investor behavior

by

Anthony Bellofatto[1], Marie-Hélène Broihanne[2], Rudy De Winne[3], Catherine D'Hondt[4], Maxime Merli[5], Hava Orkut[6], Patrick Roger[7]

August 31, 2018

[1] [1] Louvain Finance (IMMAQ), Louvain School of Management, Université catholique de Louvain - Address: Chaussée de Binche 151, 7000 Mons, Belgium - Email: anthony.bellofatto@uclouvain.be.
[2] LaRGE Research Center, EM Strasbourg Business School, University of Strasbourg. Address: 61 Avenue de la Forêt Noire, 67085 Strasbourg Cedex, France, E-mail: mhb@unistra.fr.
[3] Louvain Finance (IMMAQ), Louvain School of Management, Université catholique de Louvain - Address: Chaussée de Binche 151, 7000 Mons, Belgium - Email: rudy.dewinne@uclouvain.be.
[4] Louvain Finance (IMMAQ), Louvain School of Management, Université catholique de Louvain - Address: Chaussée de Binche 151, 7000 Mons, Belgium - Email: catherine.dhondt@uclouvain.be.
[5] LaRGE Research Center, EM Strasbourg Business School, University of Strasbourg, E-mail: merli@unistra.fr.
[6] LaRGE Research Center, EM Strasbourg Business School, University of Strasbourg, E-mail: horkut@unistra.fr.
[7] LaRGE Research Center, EM Strasbourg Business School, University of Strasbourg, E-mail: proger@unistra.fr.

Introduction

The present report aggregates the analyses and results of three empirical studies centered on the relationships between answers to MIFID questionnaires, actual trading behavior of retail investors, and behavioral differences between investors receiving advice and professional information and those neglecting this information and advice.

The MiFID1 European Directive came into force in November 2007. Its goal was to increase the level of protection of retail investors by requiring investment firms to deliver the most suitable services to their clients. Therefore, investment firms operating in the EU are committed to collect client-related information through the so-called MiFID questionnaires. The quantity and nature of information to be collected depend on the service(s) requested by investors. The European Directive defines three types of services: transmission and execution of orders, financial advice, and portfolio management. Investors who only need to execute transactions on “complex" instruments are required to fulfill the *Appropriateness test* (hereafter A-test). The A-test ensures that the investor has the necessary experience and knowledge to understand the risks involved in “complex" financial instruments before trading. Investors who ask for financial advice or portfolio management services are required to fulfill the *Suitability test* (hereafter the S-test). Assessment of suitability involves ensuring that the instruments and services offered by the intermediary meet the investor's objectives, his/her financial capacity as well as his/her knowledge and experience in financial instruments. Henceforth, MiFID requirements offer a natural field to investigate the relationships between a purposeful need for information, i.e. asking for advice, and the trading behavior of retail investors.

The three papers that form this report are entitled as follows:

1) MiFID questionnaire answers and stock market participation (Marie-Hélène Broihanne and Hava Orkut);
2) Subjective Financial Literacy and Retail Investors' Behavior (Anthony Bellofatto, Catherine D’Hondt, Rudy De Winne);
3) Appetite for Information in Mandatory Profiling of Individual Investors (Anthony Bellofatto and Marie-Hélène Broihanne).

Though the three papers were written by different authors of the team, they share common features beyond the broad topic under investigation. The three papers use two databases of retail investors from which are known, not only the answers of investors to MIFID questionnaires, but also trades and portfolios of these investors over time.[8] On an aggregate basis, our studies deal with more than 100,000 individual investors, namely 70,000 clients of a European retail bank in paper 1, and almost 60,000 clients of a Belgian online brokerage

[8] The most interesting feature of our data is that we are able to link individually the answers to MiFID questionnaires and both the trades and the portfolios of these investors.

house in papers 2 and 3.[9] Although the appropriateness and suitability tests are often parts of a single questionnaire (in particular in French banks), we were lucky enough to get data in two distinct questionnaires in our Belgian database (one for the A-test and one for the S-test) . We then discovered that approximately 50% of the Belgian sample chose not to fill in the S-test, despite the fact that this specific test gave free access to an advice tool on stocks available on the web platform (that provides both professional information and recommendations on stocks).

The papers in this report address several questions. The first paper investigates the usefulness of the MiFID questionnaire answers when it comes to understand stock market participation. The informational content of MiFID questionnaire answers is shown to be important enough so that financial institutions have to take care of the way they build their questionnaires, collect data and use them when providing their clients with suitable advice or instruments.

The second question generalizes the first one because it not only addresses stock market participation but also 1) the consistency of answers between the A-test and the S-test, and 2) the actual behavior of investors as a function of their self-reported financial literacy.[10] For that purpose, we use subjective measures of financial literacy available in the two MiFID tests of our second database. Such survey data may best capture psychological drivers affecting the individual's decision-making process. We stress that the specificity of this second database, which enabled us to distinguish A-investors and S-investors, is very interesting because it allows to compare the behavior of A-investors, who filled in only the A-test, and S-investors, who filled in both tests (A-test and S-test) and then have access to the advice tool on stocks available on the web platform. In particular, the third paper compares the performance of the two categories of investors over the period 2008-2012. In a nutshell, S-investors trade on a larger universe of stocks and hold portfolios that are more diversified than those of A-investors. They are also more active on “complex" instruments. These differences in trading behavior may explain why the S-investors earn significantly higher returns.

MiFID questionnaire answers and stock market participation

There is a large body of literature on stock market participation.[11] However, understanding retail investors' behavior requires a consistent methodology to assess individuals' attitudes towards risk. The common tools to reach this objective are either lottery choices, experiments (that present simple choices), willingness-to-pay surveys, i.e. revealed preferences, or using secondary data reflecting actual investment decisions, i.e. stated preferences. Only a few empirical studies assess retail investors' attitudes towards risk by means of a dedicated online questionnaire. The first paper of this report uses MiFID questionnaire answers to assess individuals' attitudes. In fact, MiFID requires investment firms operating in the EU to get a thorough knowledge of their clients to offer advice and financial

[9] In fact, papers 2 and 3 deal with subsamples of the Belgian database because our analyses are focused on the “MiFID period”, which starts at the end of 2007, while the data on investors’ accounts cover the 2003-2012 period.

[10] Lusardi and her co-authors define financial literacy as the ability to process economic information and make informed decisions about financial planning, wealth accumulation, debt and pensions.

[11] As the detailed references are given in the papers, we do not quote any paper in this introduction of the report.

products that are perfectly suited to the clients' situation. We especially focus our attention on the self-assessed risk tolerance and attitudes towards losses, and test their impact on stock investment decision.

The use of the MiFID questionnaire answers gives useful insights since retail investors have a strong tendency to give relevant answers in order to get well-suited advice from their financial advisers. Besides, matching questionnaire answers with banking records ensures high data reliability. Finally, the mandatory data collection through MiFID questionnaires allows us to analyze a significant sample wherein male and female investors are equally represented. In addition, our sample is representative of the whole French population regarding socio-demographics and wealth level, including income and credit. Therefore, our results are exempt from any selection bias.

Our results show that the MiFID questionnaire answers are strong indicators of stock market participation. Indeed, retail investors characterized by a high-risk tolerance level and a low sensitivity towards losses are more likely to hold stocks. More precisely, the impact of risk tolerance on stockholding decision is greater than that of attitude towards losses. We also find that natives and those living in the Paris region are more likely to hold stocks. Furthermore, retail investors who opt for the separation regime are more prone to invest in stocks. The same pattern is observed for self-employed investors. These results show that financial independence promotes stock market participation. Not surprisingly, the more credit-constrained investors are less likely to hold stocks.

Finally, the holding of other investment vehicles, such as unit-linked life insurance and/or retirement plans, is positively linked to the decision to hold stocks. Our findings are still valid on the specific subsample of investors who are familiar both with the MiFID questionnaire content and with the stock market. In particular, a high-level of financial literacy is strongly associated with stockholding.

Our results contribute to the ongoing debate across professionals, regulators and academics about the usefulness of MiFID questionnaires. As already pointed out in recent studies, we demonstrate that a quantitative measurement of risk-taking preferences is necessary for ensuring that investment service providers offer well-suited advice and financial products to their clients. Furthermore, this paper gives insights to improve MiFID questionnaires. As there is no regulatory constraint on the questionnaire content, investment firms are free to establish their own questionnaires (provided that they respect some general guidelines). We claim that attention should be devoted to the length of the questionnaire. Even if each question brings additional information, fatigue or mistrust should be considered. Besides, questionnaire length may also influence the number of unreported answers, which may restrict the ability of investment firms to offer suited advice to their clients. For these reasons, we suggest to build questionnaires as short as possible.

Subjective Financial Literacy and Retail Investors' Behavior

The second paper investigates the relationship between trading behavior and subjective financial literacy, i.e. the level of financial knowledge and experience reported in the MiFID tests by retail investors. In this paper, we use data about MiFID tests that were conducted online and answers are then self-reported decisions the investors make on their own. The advantage is that the answers are not affected by any conversation with a broker or a financial advisor. However, online tests - like online trading activities in general - have also a drawback: they make investors "do-it-yourselfers" in an information-rich environment, thereby bolstering their overconfidence due to an illusion of both knowledge and control.

For the purpose of this paper, we focus on a subsample of 20,285 retail investors of our second database to investigate the behavior of investors over the period 2003-2012. We first check the consistency between subjective financial literacy reported in the S-test and A-test. In a second step, we check the consistency between subjective financial literacy and trading behavior characterized along three different aspects: experience and familiarity with financial markets, diversification and performance.

Our main findings show an overall consistency across investors' answers: investors who report a high literacy in the A-test are much more likely to also report a high literacy in the S-test. As mentioned before, we also find that subjective literacy partly explains cross-sectional variations in investors' behavior. Investors who self-report a high level of financial literacy trade more on both stocks and complex instruments and they are less prone to the disposition effect, which is a result consistent with a deeper experience of financial markets. Although highly literate investors trade a larger universe of stocks, the stock portfolios they hold are less diversified. In fact, though they concentrate their stock portfolios on a small set of securities, they achieve a global portfolio diversification through investing in funds (that are already diversified by nature). Finally, investors with a higher level of subjective financial literacy display higher gross and net returns and higher Sharpe ratios as well. These results hold even when we control for gender, age, portfolio value, trading experience and education.

All in all, our findings support consistency between subjective literacy and actual trading behavior. Retail investors are overall consistent when reporting their financial literacy online. More importantly, this piece of information provided by the investors themselves could help better understand and characterize their actual trading behavior. Such results are relevant for both policy-making and understanding retail investors' behavior. Subjective literacy reported in the MiFID tests is informative to characterize retail investors and hence deserve more attention in that perspective. This empirical evidence is meaningful for investment firms that are committed to administer the MiFID tests in the EU. This paper could also provide insights for regulatory purposes, since we show that subjective financial literacy reported online does correlate with actual trading behavior.

Appetite for Information in Mandatory Profiling of Individual Investors

The third paper investigates the trading activity of 14,155 retail investors over the 2008-2012 period (second database). This subsample is exclusively made of investors who started trading at the Belgian brokerage house after the MiFID directive came into force. Since this online brokerage house does not offer portfolio management services during the period under scrutiny, investors in our sample have either fulfilled the A-test to execute transactions or both the A- and S-test to have access to a professional advice tool on stocks. S-investors have revealed a willingness to access a service that goes beyond order execution only ("premium service"). However, the only cost of this supplementary service is the time needed to fill in the S-test. A by-product of the S-test is that S-investors reveal what we call "appetite for information". On the contrary, A-investors voluntarily neglect a free access to professional recommendations, suggesting a tendency to behave more intuitively.

The results of our empirical work show that S-investors trade on a larger stock universe, hold better diversified portfolios, and are more active on “complex" instruments. By contrast, A-investors concentrate their trades on a lower number of stocks, execute more day-trades and roundtrips on this set of stocks and are less attracted by non-stock instruments. These trading behavior differences may explain why the S-investors earn significantly higher returns. This finding holds even under a random matching procedure that controls for socio-demographic data, financial experience, education and other various survey answers.

This paper contributes to two strands of literature: (a) literature on the relationship between trading behavior and information acquisition and (b) literature on the relationship between trading behavior and personality traits. While the vast majority of papers are only descriptive, our analysis deals with actual trading records. Furthermore, unlike preceding papers, we do not focus on trading frequency only but analyze trading behavior in a broader sense, including trading performance.

Our findings have also implications for regulators (FED, ESMA, AMF) and investment firms. They suggest that investors' behavior is consistent with attitude towards financial information. In line with their choice to neglect a free access to an information tool, the A-investors trade more intuitively while investors displaying “appetite for information” tend to behave more in line with the traditional Finance theory.

MiFID questionnaire answers and stock market participation

Marie-Hélène Broihanne[a] and Hava Orkut[b]

April 4, 2018

Abstract

The Markets in Financial Instruments Directive (MiFID) aims at strengthening the protection of investors by requiring investment service providers to submit a questionnaire to their clients. In this paper, we combine MiFID questionnaire answers and banking records of more than 70,000 retail clients of a big European retail bank. We demonstrate that *MiFID indicators*, i.e. self-assessed risk tolerance and attitudes towards losses, explain stock investment decision while controlling for gender, age and income. We show that MiFID indicators exhibit greater magnitude effects than those of stock investment classical determinants. We also demonstrate that investor country of birth, residency, matrimonial regime and holding of other risky financial products, such as unit-linked life insurance and retirement plans, are important drivers of stock market participation. Our results are consistent with prior studies on the determinants of stock market participation and are robust to robustness checks.

Keywords : Stock investment, MiFID questionnaire, risk tolerance, attitudes towards losses

JEL Classification : G02, G11, G28

LaRGE Research Center, EM Strasbourg Business School, University of Strasbourg

Address: 61 Avenue de la Forêt Noire, 67085 Strasbourg Cedex, France

(a) E-mail: mhb@unistra.fr / Corresponding author

(b) E-mail: horkut@unistra.fr

1 Introduction

Understanding investment behavior of retail investors is important for asset pricing under well-known investment mistakes such as under-diversification[1] and non-participation[2] in the financial and insurance markets. Numerous empirical works have examined retail investors' decisions and their consequences on asset prices (e.g. Hirshleifer, 2001, Brown and Cliff, 2005 and Baker and Wurgler, 2006, 2007). For example, investor sentiment (Baker and Wurgler, 2007) drives asset pricing (Yang and Li, 2013, Blau, 2017 and Ryu et al., 2017) and is positively correlated with assets mispricing (Brown and Cliff, 2005 and Chang et al., 2015). Among others, retail investors' overconfidence (Daniel and Hirshleifer, 2015) or the disposition effect (Chang et al., 2016) also impact asset prices. However, classical asset pricing theories, such as CAPM (Sharpe, 1964, Lintner, 1965 and Mossin, 1966), make the usual assumption that investors are rational and risk averse. Therefore, understanding retail investors' behavior makes necessary the use of a consistent methodology to assess attitudes towards risk.

Attitudes towards risk are usually assessed either by using hypothetical lottery choices (Holt and Laury, 2002 and Booij and Van de Kuilen, 2009) or willingness-to-pay surveys (Cummings et al., 1986 and Mitchell and Carson, 1989), i.e. *revealed preferences*, or, by using secondary data reflecting actual investment decisions (Barber and Odean, 2001), i.e. *stated preferences*. Only a few empirical studies assess retail investors' attitudes towards risk by using a dedicated online questionnaire (Dorn and Huberman, 2005 and Hoffmann et al., 2013, 2015). In these works, assessed attitudes towards risk are combined with brokerage records to study investors' trading activity.

In this paper, we employ an alternative approach by using the Markets in Financial Instruments Directive[3] (MiFID) questionnaire answers to assess individuals' attitudes. This directive requires investment service providers to get a thorough knowledge of their clients thanks to the MiFID questionnaire[4] to offer advice and financial products perfectly suited to their situation, which is not the major concern of surveys/questionnaires used by prior behavioral finance works. Overall, the MiFID questionnaire answers allow highlighting clients' declared characteristics, needs and preferences and establishing their risk profile.

In this paper, we combine the MiFID questionnaire answers of more than 70,000 retail clients of a large European retail bank to their banking records to explain stock market participation. There is a large body of literature on stock market participation (Peress, 2005, Bogan, 2008, Guiso et al., 2008, Grinblatt et al., 2011, Bonaparte and Kumar, 2013, Antoniou et al., 2015, Fischer and Jensen, 2015 and Cronqvist et al.,

[1] See Lease et al. (1974), Odean (1999), Mitton and Vorkink (2007), Kumar (2007) and Goetzmann and Kumar (2008).

[2] See Mankiw and Zeldes (1991) and Poterba and Samwick (1995).

[3] Implemented in 2007, MiFID I (2004/39/EC) gathers 31 member states of the European Economic area (28 European member states and 3 other states: Iceland, Norway and Liechtenstein). It replaces the Investment Services Directive (ISD) adopted in 1993. MiFID I aims to improve the competitiveness of European financial markets and to ensure a harmonized protection to individuals according to their level of financial knowledge. From January 2018, MiFID II (2014/65/UE) replaces MiFID I (2004/39/EC) that we consider in this paper. MiFID II aims to strengthen the transparency, the efficiency of financial markets but also the protection of investors. Note that MiFID questionnaire is only imposed to the MiFID member states whereas it is not used in the US.

[4] Note that the directive does not impose a standard questionnaire. Each bank is free to prepare and organize its own questionnaire.

2016). Like Hong et al. (2004), Fan and Xiao (2006) and Kaustia and Torstila (2011), we do not focus on quantitative differences in stock holding but only on the decision to hold.

Our results show that stock market participation is driven by risk tolerance and attitudes towards losses, both being assessed in the MiFID questionnaire. Indeed, retail clients with a high risk tolerance level and those with a low sensitivity towards losses are more likely to hold stocks. We show that the impact of risk tolerance is greater on stockholding decision than that of attitudes towards losses. We also find that natives and those living in the Paris region are more likely to hold stocks. Further, retail clients opting for the separation regime are more prone to invest in stocks. The same pattern is observed for self-employed retail clients. We conclude that financial independence promotes stock market participation. As for wealth and patrimony indicators, we find that high credit-constrained retail clients are less likely to own stocks. Finally, holding other investment vehicles, such as unit-linked life insurance and/or retirement plans, has a positive impact on the decision to hold stocks.

Our findings bring new insights on stock market participation for the four following reasons.

First of all, the use of the MiFID questionnaire answers gives useful insights since retail clients must give relevant answers in order to get well-suited advice from their financial adviser. The informativeness of the MiFID questionnaire to characterize retail investors' behavior has been demonstrated by Bellofatto et al. (2015) for Belgium retail investors. In our paper, we focus on the self-assessed risk tolerance and attitudes towards losses, both being seldom analyzed together in behavioral finance works. Besides, matching questionnaire answers with banking records ensures high data reliability.

Second, the mandatory data collection through the MiFID questionnaire allows to analyze a significant sample where male and female retail clients are equally represented. Further, our sample is representative of the whole French population regarding socio-demographics and wealth level, including income and credit. Therefore, our results are exempt of any selection bias and any stockholding differences between, for example men and women (Barber and Odean, 2001, Dwyer et al., 2002 and Agnew et al., 2003), are controlled for other covariates in our data sample.

Third, our paper is the first to combine declared information altogether with real investment decisions for studying stock market participation. Indeed, Hong et al. (2004), Fan and Xiao (2006), Balloch et al. (2014), Georgarakos and Inderst (2014) and Liang and Guo (2015) only use surveys for determining stock market participation determinants. We introduce a variety of variables, some of which are rarely studied, like the country of birth and residency, or never yet studied, like the matrimonial regime choice. For the matrimonial regime, attention is paid to the separation regime which allows spouses to be financially independent from each other. Overall, we can compare the magnitude effects of these variables to those of classical ones considered as key drivers of investment decisions such as gender, age and income.

Finally, MiFID data are for the first time introduced for explaining stock market participation[5] whereas

[5]Cumming et al. (2011) and Aitken et al. (2015) focus on MiFID and exchange trading rules.

they were previously used for analyzing the disposition effect in Belgium (Bellofatto et al., 2014), the risk profile in Italy (Mazzoli and Marinelli, 2014) and investor sentiment in Belgium (D'Hondt and Roger, 2018). By the way, this study contributes to the literature on French retail investors that already addressed the disposition effect (Boolell-Gunesh et al. 2009, 2012), diversification (Broihanne et al., 2016), herding behavior (Merli and Roger, 2013), the market sentiment index (Roger, 2014), market performance (Magron 2012, 2014), the impact of retail investors on market volatility (Foucault et al., 2011), the influence of investor attention on the French stock market activity and trading (Aouadi et al., 2013), the relationship between financial literacy and portfolio rebalancing (Bianchi, 2018) or the phenomenon of overreaction and underreaction on French stock market (Siwar, 2011). Stockholding in France has also been documented by Arrondel et al. (2015) who demonstrate the positive impact of basic financial literacy on stock market participation. In their paper, data come from a household survey, whereas in ours, stockholding is filled out in the banking records.

Our paper is organized as follows. Section 2 describes our datasets. Section 3 displays empirical results. Section 4 is dedicated to robustness checks. Section 5 concludes.

2 Data and descriptive statistics

In this study, we use two datasets provided by a large European retail bank. The first dataset (Dataset 1) contains MiFID questionnaire answers of more than 70,000 retail clients over the period 2007-2015. The second dataset (Dataset 2) includes banking records of these retail clients on the 07/31/2015. Therefore, our datasets differ from the ones of studies where brokerage house data of retail investors are used, in France (Boolell-Gunesh et al., 2009, 2012) or in other European countries (Belgium (Bellofatto et al., 2014), Germany (Weber et al., 2013), Netherlands (Hoffmann et al., 2015) and UK (Richards et al., 2017)) but also in the non-European countries such as in China (Chen et al., 2007 and Feng and Seasholes, 2005, 2008) and the US (Barber and Odean, 2001 and Korniotis and Kumar, 2011).

A variety of variables is used throughout this study. We classify them into 3 Panels: MiFID indicators (Panel A and Section 2.1), socio-demographic indicators (Panel B and Section 2.2) and wealth and patrimony indicators (Panel C and Section 2.2). We exclude individuals aged under 18 years old from the analysis as in Bauer et al. (2009) and Hoffmann et al. (2013, 2015).

Table 1 defines variables. Table 2 displays descriptive statistics.

In our sample, 11.05% of retail clients directly or indirectly[6] hold stocks as of the 07/31/2015[7]. Such a

[6] In France, stocks are indirectly held *via* "Equity Saving Plan" (*Plan Epargne en Actions* or *PEA* in French) which offers fiscal advantage to their holders. Specifically, capital gains are tax free as the PEA account has been held at least 5 years ago. French specificities have been documented by Boolell-Gunesh et al. (2009). In our study, when stockholding is indirect, "Stocks" is set one if the retail client holds at least one "active" PEA account as of the 07/31/2015, i.e. the quantity of stocks held differs from zero, and zero otherwise (if there is only cash in the saving plan for example).

[7] According to SoFia survey of TNS Sofres, the number of individuals holding financial assets in France decreased from 20% to 11% and stockholding rate decreased from 15.9% to 8.1% between 2009 and 2015.

low stock market participation has already been documented in the US (Mankiw and Zeldes, 1991, Haliassos and Bertaut, 1995 and Poterba and Samwick, 1995), the UK (Attanasio et al., 2002) and France (Arrondel et al., 2015). Information costs (Mankiw and Zeldes, 1991), fixed participation costs (Vissing-Jorgensen, 2002), social interaction (Hong et al., 2004), the lack of financial awareness (Guiso and Jappelli, 2005), internet access (Bogan, 2008), the lack of trust (Guiso et al., 2008), IQ scores (Grinblatt et al., 2011), political preferences (Kaustia and Torstila, 2011), redistributive tax system (Fischer and Jensen, 2015) and stock market image (Dobni and Racine, 2015, 2016) may explain limited stock market participation.

Table 1 – Variable description

Variables	Definitions
Dependent variable	
Stocks	Dummy variable coded 1 if the client directly or indirectly holds stocks as of the 07/31/2015 and 0 otherwise.
Independent variables	
	Panel A: MiFID indicators
Risk tolerance	Self-assessed risk tolerance level ranged from 0 (no risk tolerance) to 2 (high risk tolerance).
Attitudes twd losses	Self-assessed attitudes towards losses during a hypothetical downturn for a portfolio securities. Attitudes are ranged from 1 to 4: 1 (selling the entire portfolio), 2 (selling a part of the portfolio), 3 (waiting until portfolio value increases) and 4 (taking advantage of lower price to invest again).
	Panel B: Socio-demographic indicators
Gender	Dummy variable coded 1 for males and 0 for females.
Age	Age of the client as of the 07/31/2015 (in years).
Native	Dummy variable coded 1 if the client is native of the country and 0 otherwise.
Paris	Dummy variable coded 1 if the client lives in and close to the Paris region and 0 otherwise.
Matrimonial	Dummy variable coded 1 if the client has chosen the separation regime and 0 otherwise.
Self-employed	Dummy variable coded 1 if the client perceives directly his/her income from his/her own professional activity and 0 otherwise.
Salaried	Dummy variable coded 1 if the client has a wage or salary from an employer and 0 otherwise.
Retired	Dummy variable coded 1 if the client is retired and 0 otherwise.
No occupation	Dummy variable coded 1 if the client has no occupation (e.g. students or any professional activity) and 0 otherwise.
	Panel C: Wealth and patrimony indicators
Income	Net monthly income (in euros).
Credit	Credit amount remaining to be reimbursed (in euros).
UL life insurance	Dummy variable coded 1 if the client holds unit-linked life insurance products as of the 07/31/2015 and 0 otherwise.
Retirement	Dummy variable coded 1 if the client holds retirement plans as of the 07/31/2015 and 0 otherwise.

Table 1 describes all variables. Independent variables are classified into three panels: Panel A: MiFID indicators; Panel B: Socio-demographic indicators and Panel C: Wealth and patrimony indicators.

Table 2 – Descriptive statistics

		N	$\overline{X}$/ %	std	min	max
Retail clients		77,365	100%	-	-	-
Dependent variable						
Stocks		77,365	11.05%	-	-	-
Independent variables						
Panel A : MiFID indicators						
Risk tolerance		71,461	0.32	0.50	0	2
0			69.35%	-	-	-
1			28.90%	-	-	-
2			1.75%	-	-	-
Attitudes twd losses		71,745	2.71	0.78	1	4
1			14.29%	-	-	-
2			6.24%	-	-	-
3			73.93%	-	-	-
4			5.54%	-	-	-
Panel B : Socio-demographic indicators						
Gender		77,365	51.24%	-	-	-
Age		77,365	47.97	17.55	18	105
Native		77,365	84.59%	-	-	-
Paris		77,365	12.26%	-	-	-
Matrimonial		77,365	10.30%	-	-	-
Self-employed		77,365	12.61%	-	-	-
Salaried		77,365	55.36%	-	-	-
Retired		77,365	15.59%	-	-	-
No occupation		77,365	16.44%	-	-	-
Panel C : Wealth and patrimony indicators						
Income		77,365	2,418.07	2,192.97	0	10,000
			1.90	1.11	0	5
INCOME BRACKETS :	CODES :					
0	0		7.28%$^{(0)}$	-	-	-
<1,500	750		31.62%$^{(1)}$	-	-	-
1,500-3,000	2,250		36.67%$^{(2)}$	-	-	-
3,000-5,000	4,000		15.32%$^{(3)}$	-	-	-
5,000-10,000	7,500		6.72%$^{(4)}$	-	-	-
>10,000	10,000		2.39%$^{(5)}$	-	-	-
Credit		77,365	28,668.91	38,960.65	0	100,000
			1.04	1.18	0	3
CREDIT BRACKETS :	CODES :					
0	0		50.08%$^{(0)}$	-	-	-
<10,000	5,000		13.51%$^{(1)}$	-	-	-
10,000-100,000	55,000		18.70%$^{(2)}$	-	-	-
>100,000	100,000		17.71%$^{(3)}$	-	-	-
UL life insurance		77,365	16.83%	-	-	-
Retirement		77,365	1.37%	-	-	-

Table 2 presents descriptive statistics of the variables in Panels A, B and C. The first column reports variable names. For each variable, the second column reports the number of retail clients (N) for which the data is available. The third column reports the percentage (%) of retail clients for which the corresponding variable is equal to one for binary variables and the mean ($\overline{X}$) for continuous variables. For MiFID dummy variables, "Income" and "Credit", we provide in addition the percentage corresponding to each of the modalities indicated in parentheses and superscript. The fourth column reports the standard deviation (std). The fifth and sixth columns indicate minimum (min) and maximum (max) values.

2.1 MiFID questionnaire answers

The MiFID questionnaire of the bank includes six sections dealing with socio-demographics, income, patrimony, credit, savings capacity and investment goals, respectively. The latter section is composed of four subsections dealing with the main investment goals, risk tolerance, experience with financial products and attitudes towards losses during a hypothetical downturn.

The questionnaire was administered a maximum of three times to retail clients between 2007 and 2015[8]. The questionnaire was administered the first time to any client who subscribed any financial instrument after 2007. The second questionnaire was administered three years after the first one. The third questionnaire was administered to retail clients subscribing any financial instrument after a second one, or, three years after the second one. Answers provided by retail clients to whom the questionnaire was administered at least twice are stable over time[9]. For that reason, we only focus on the most recent MiFID questionnaire answers of these retail clients. Besides, the most recent questionnaire answers are always extracted at a date which is the closest to and occurred before 07/31/2015, i.e. the date of banking records extraction.

In the questionnaire, we pay attention to retail clients' self-assessed risk tolerance ("Risk tolerance") and attitudes towards losses ("Attitudes twd losses"). Risk tolerance corresponds to the level of risk that a retail client is accepting to bear. Attitudes towards losses correspond to the behavior that a retail client would exhibit if his/her portfolio value would decrease by 15%[10]. These variables are at the most empirical proxies for the concepts of risk aversion and loss aversion[11] usually found in the decision-making literature. In the questionnaire, risk tolerance and attitudes towards losses have not been reported by 7.63% and 7.26% of retail clients, respectively. Both variables have been shown to impact trading activity and portfolio composition. For example, Hoffmann et al. (2015) find that investors with higher levels of and upward revisions in risk tolerance are more likely to trade and hold riskier portfolios. Using household survey data, Dimmock and Kouwenberg (2010) find that higher loss aversion reduces the probability of participating in equity markets.

Looking at "Risk tolerance" (Table 2, Panel A), we notice that many retail clients are not risk tolerant at all (69%). This large proportion may explain the low level of stock market participation. About 31% of retail clients consider themselves risk tolerant. Likewise, Hong et al. (2004) find that 32.53% of US households are risk tolerant in their sample[12].

[8]Prepared by the bank, the questionnaire was the same over the period 2007-2015. It satisfies both MiFID requirements and anti-money-laundering standards. Retail clients have answered the questionnaire with a financial adviser of the bank. About 70% of them use the checking account opened within the bank on a daily basis.

[9]We compute agreement rates for answers between two successive questionnaires and find that, on average, 90.69% of retail clients report the same self-assessed attitudes over time.

[10]The percentage of loss was freely chosen by the bank and reflects the volatility of the stock market which is around 15-20% on an annual basis. Using the MiFID questionnaire of a Belgium online bank, Bellofatto et al. (2014) study a similar question in which the percentage of loss is 20%.

[11]Tversky and Kahneman (1992) introduce the Cumulative Prospect Theory (CPT) to illustrate the asymmetry existing between gains and losses. According to their experimental work, they estimate a loss aversion coefficient (λ) which allows to quantify the sensitivity towards losses. When λ is higher than 1, individuals exhibit loss aversion. They show that individuals are 2.25 times more sensitive towards losses than gains.

[12]However, Hong et al. (2004) use a dummy variable indicating risk tolerance.

As for "Attitudes twd losses", there are four modalities which selling rate decreases from the first to the last modality: (1) selling the entire portfolio, (2) selling a part of the portfolio, (3) waiting until portfolio value increases and (4) taking advantage of lower price to invest again. In our sample, a large number of retail clients (about 74%) would wait until their portfolio value increases in case of a downturn. About 20% of them would sell the entire or a part of their portfolio. Only 5.54% of retail clients would purchase additional financial securities during the downturn.

2.2 Banking records

In Panel B, we first point out that men represent 51.24% of the sample. Gender parity is seldom observed in behavioral finance works. Indeed, male retail investors usually represent about 80% of the sample (Boolell-Gunesh et al. (2009) in France, Bellofatto et al. (2014) in Belgium, Weber and Welfens (2007) in Germany, Bauer et al. (2009) in the Netherlands, Richards et al. (2017) in the UK, Grinblatt and Keloharju (2009)[13] in Finland and Barber and Odean (2001) in the US). However, our sample is similar to that analyzed in East Asia (Feng and Seasholes, 2008) due to a high participation rate of Chinese women on financial markets compared to other countries (Chen et al., 2004 and Feng and Seasholes, 2005).

Previous research has identified gender differences in financial decision making. We assume that male retail clients are more likely to invest in stocks than their female counterparts. Indeed, women are less likely to participate in the stock market than men (Haliassos and Bertaut, 1995, Van Rooij et al., 2011 and Almenberg and Dreber, 2015). They hold less risky assets (Riley and Chow, 1992, Hinz et al., 1997, Bernasek and Shwiff, 2001, Dwyer et al., 2002, Agnew et al., 2003, Charness and Gneezy, 2012 and Georgarakos and Inderst, 2014) and are less risk seeking (Powell and Ansic, 1997, Jianakoplos and Bernasek, 1998, Halek and Eisenhauer, 2001, Booij and Van de Kuilen, 2009 and Booth and Nolen, 2012) than men. Women allocate a smaller percentage of their financial assets to stocks than to bonds (Bajtelsmit and VanDerhei, 1997 and Bajtelsmit et al., 1999). They have a more conservative approach as they consider financial markets as being riskier than men (Jacobsen et al., 2014). Barber and Odean (2001) show that men trade stocks 45% more often than women.

On average, retail clients are 48 years old. Our sample is slightly older than that of Feng and Seasholes (2005) in China (about 35 years) and slightly younger than those of Hallahan et al. (2004) in Australia (the largest number of individuals belongs to the age bracket 51-60 years), Dhar and Zhu (2006) in the US (50 years) and Van Rooij et al. (2011) in the Netherlands (about 51 years).

Working on asset allocation decisions of US households, Ackert et al. (2002) show that age has an impact on the mix of risky assets. Indeed, young households prefer investing more in stocks than in bonds. Guiso et al. (2008) indicate that the probability of direct participation in the stock market decreases with age. Bodie and Crane (1997) find that the proportion of risky assets held by individuals decreases with age. Bakshi

[13] Working on a dataset from Finland, Grinblatt and Keloharju (2009) only focus on men enlisted into mandatory military service.

and Chen (1994) show that risk aversion increases with age. Recently, Brooks et al. (2018) confirmed the negative relationship between age and risk tolerance by showing that risk tolerance declines at an increasing, albeit slow, rate with age.

However, the impact of age on individuals' investment decisions is not always clear-cut. Differences in findings may be attributed to the methodologies employed (laboratory experiments, surveys or portfolio holdings) and sample characteristics (households, investors). For example, Shum and Faig (2006) study asset allocation of US households and show that the decision to own stocks is positively correlated with age. Their result is in line with that of Balloch et al. (2014). Wang and Hanna (1997) show that the proportion of net wealth invested in risky assets increases with age. An important point is that age and risk aversion are linked together. For example, Grable (2000) finds that risk tolerance increases with age on a sample of faculty and staff working at a large university. Other studies indicate that the relationship between age and risk aversion is not linear. By deriving relative risk aversion indexes from actual asset allocations of the US population, Riley and Chow (1992) find a U-shaped relationship, i.e. risk aversion decreases with age then increases after 65 years old. By using psychometrically validated survey, Faff et al. (2008) show that young and older individuals are more risk tolerant compared to those who are middle aged[14]. For these reasons, testing the impact of age on stockholding is an interesting question in our data.

Two geographical criteria are analyzed. We know whether retail clients were born in the country ("Native") but also whether they are living in the Paris region ("Paris"). This latter variable allows to test the impact of the biggest region of the country, in economic and size terms, on stock investment decision. In our sample, about 85% of retail clients are French native-born and 12% of them live in or close to the Paris region[15]. We assume that natives are more likely to invest in stocks like Osili and Paulson (2004) and Chatterjee (2009). Demographic breakdown is also analyzed by Tekce and Yilmaz (2015) for explaining Turkish retail investors' behavior on stock market. About half of their sample live in the most developed region containing the biggest Turkish city, Istanbul, and about 17% live in the region where the capital, Ankara, takes place. In our study, "Paris" allows capturing these two specificities. First, regarding the gross domestic product (GDP) per region, the Paris region displays the highest economic performance of the whole country. Second, it includes Paris which is both the capital and the largest French city. Previously, Arrondel et al. (2010) find that households living in Paris are more likely to invest in risky assets. Then, we expect that retail clients living in or close to the capital are more likely to hold stocks.

Matrimonial regime choice is introduced for the first time in this study. Complementing marital status, matrimonial regime structures patrimony allocation rules within spouses during the marriage but also after its breakdown (for divorced or widowed individuals). Among the different French matrimonial regimes[16],

[14]Faff et al. (2008) analyze the link between financial risk tolerance and risk aversion and show that they are strongly aligned for explaining decision-making under uncertainty.

[15]The Paris region (or *Île-de-France* in French) represents an administrative region of France concentrating about one-third of national wealth, according to INSEE (the French national statistics bureau) in 2013. 11.6% of the French population were born in a foreign country and about 18.8% of them live in Paris region (INSEE, 2014). Our sample is then similar to the French demography.

[16]French matrimonial regimes are divided into two categories: community and separation regimes. Community regimes are based

we pay attention to the separation of property regime ("Matrimonial") which implies no joint-ownership between spouses. Under this regime, each spouse is free to manage his or her own goods and is liable for any debt incurred by him or her before and after the marriage. We then consider "Matrimonial" as a proxy for patrimony protection desirability of retail clients. In our sample, 10% of retail clients have chosen the separation regime[17]. We expect that the separation regime increases the likelihood of stock investment. We consider that, like marriage (Bertocchi et al., 2011), it represents another kind of safe asset in a portfolio choice framework. Actually, married households are more likely to invest in risky assets than single ones (Agnew et al., 2003 and Bertocchi et al., 2011). Indeed, Agnew et al. (2003) analyze nearly 7,000 retirement accounts and find that stock allocation is higher among married investors than among their single counterparts. This finding is consistent with that of Grable (2000) who shows that married individuals are more risk tolerant. Nevertheless, Grable and Joo (2004) find the reverse evidence by analyzing a sample of faculty and staff from two large universities.

Several professional categories are available in Dataset 2. We group and recode them as 4 dummy variables: "Self-employed", "Salaried", "Retired" and "No occupation". The distinction between self-employed and salaried has already been studied by Maccrimmon and Wehrung (1986) and Dorn and Sengmueller (2009)[18]. According to Maccrimmon and Wehrung (1986), individuals perceiving their incomes directly from their own activity are willing to take more risk compared to those having straight salary work or wage from an employer. They are then willing to choose riskier investments than salaried individuals do. In a similar vein, Georgarakos and Inderst (2014) show that self-employed are more likely to own stocks than employed. However, Heaton and Lucas (2000) show that entrepreneurs represent a significant part of stock holders but they invest less wealth in stocks compared to other similarly wealthy households. For these reasons, testing the impact of financial independence on stockholding is interesting in our sample where salaried represent more than half of retail clients. About 13% of retail clients are self-employed. This proportion is like those indicated by Sung and Hanna (1996) and Dorn and Sengmueller (2009). 15.59% of retail clients are retired. Likewise, Van Rooij et al. (2011) report that 18.4% of Dutch households are retired. Finally, 16.44% of our sample exercise no professional activity.

In Panel C, we analyze the net monthly income of retail clients. Our approach differs from the one of other studies. For example, to measure wealth, Hong et al. (2004) use the value of all assets except for non-retirement stockholdings. Guiso et al. (2008) combine financial wealth and income. Dorn and Sengmueller (2009) use the total net worth including all financial assets and real estate. Cho (2014) assumes that wealth consists only of financial assets and housing wealth. In this paper, we do not focus on the

on the notion of common goods. Separation regime implies no joint-ownership between spouses. In Europe, community regime is the default option (e.g. Belgium, France, Italy and Luxembourg) whereas separation regime is fixed by default in other countries (e.g. Germany, England, and Greece). In the US, the legal matrimonial regime differs from one state to another one.

[17] According to INSEE, 10% of married individuals live under the separation regime in 2010.

[18] Other studies use 3 categories, i.e. professional, non-professional and retired or non-employed categories (Dhar and Zhu, 2006 and Goetzmann and Kumar, 2008), or distinguish finance-related jobs from other activities (Grinblatt and Keloharju, 2009 and Fuertes et al., 2014).

whole wealth, including real estate and financial assets, and analyze separately all components of wealth, i.e. net monthly income, credit amount (including consumer and real estate credits), unit-linked life insurance and retirement plan holdings. Actually, distinguishing income from the other wealth components allows providing a detailed description of individuals' financial situation (Kumar, 2009 and Grinblatt et al., 2011). In our sample, retail clients perceive, on average, €2,418 per month[19] and half of them have still credit to reimburse[20]. To preserve a significant sample size, the net monthly income and credit amount have been extracted from Dataset 1. Different income and credit brackets are reported in the questionnaire (see Table 2). They are attributed to numerical modalities. We use "monetary" codes which correspond to the midpoint values of income and credit brackets (the lower bound being used for the last bracket). This treatment allows us considering "Income" and "Credit" as continuous variables. Income and credit impact stock investment decisions. Shum and Faig (2006) find that the decision to own stocks is positively associated with financial net worth and labour income. Likewise, Agnew et al. (2003) find that equity allocations are higher among investors with higher income. Barber and Odean (2001) indicate that individuals having a higher income are more prone to accept market risk. Grinblatt et al. (2011) show that Finish investors belonging to the highest income decile are much more likely to participate in the stock market than the other deciles. Christelis et al. (2010) and Balloch et al. (2014) find that raising income increases the probability to invest in stocks. As for credit, Guiso et al. (1996), Fratantoni (1998) and Cardak and Wilkins (2009) show that credit-constrained households are less likely to hold risky assets. In a similar vein, investors who are less liquidity constrained are more likely to invest in stocks (Guiso and Sodini, 2013). Borrowing constraints are responsible for the limited equity investment of young consumers (Constantinides et al., 2002). Furthermore, Becker and Shabani (2010) show that households with mortgage debt are less likely to hold stocks than households with no mortgage debt. We expect to find similar results.

In our sample[21], 16.83% of retail clients invest in unit-linked life insurance[22], like in the US (Bricker et al., 2014). Arrondel et al. (2010) report a similar proportion by using survey data on French households. Only 1.37% of retail clients hold retirement plans[23] whereas about a half of US households hold such accounts (Bricker et al., 2014). The huge difference is explained by different systems in France (pay-as-you-go) and in the US (defined benefit plans and defined contribution plans).

[19]This amount is close to the average in the whole French population. Indeed, the net monthly income reported by INSEE is about €2,225 (in 2014).

[20]By only focusing on indebted retail clients, the credit amount still to repay is about €57,426. This amount is also consistent with national statistics. Indeed, INSEE reports that 47% of French households have credit to repay which is, on average, about €61,900 (in 2010).

[21]In this study, we do not analyze bonds due to the weak proportion of bondholders (lower than 1%).

[22]In France, two life insurance contracts exist: products in euros and unit-linked products. Products in euros do not generate any capital risk. As for unit-linked life insurance products, they are investment vehicles allowing to invest in different assets such as stocks, bonds or funds. Consequently, they depend on financial markets' performance.

[23]"Retirement" includes "Popular Retirement Savings Plan" *(Plan Epargne Retraite Populaire* or *PERP* in French) and "Retirement Savings Plan" *(Plan Epargne Retraite* or *PER*). Both are contracts in which the customer indirectly invest amounts on financial supports such as stocks, mutual funds, etc. They imply financial market participation.

3 Empirical results

In this paper, we use a binary logit model wherein the decision to participate in the stock market is coded 1 (Fan and Xiao, 2006, Kaustia and Torstila, 2011 and Halko et al., 2012). Our approach thus differs from that of Korniotis and Kumar (2011), who focus on the quantity of stocks held, and from that of Wachter and Yogo (2010), who study the share of financial wealth invested in stocks.

As risk tolerance and attitudes towards losses (Panel A) are correlated (Spearman rank correlation coefficient r_{sp}= 0.25 and significant at 1%), we separately analyze them.

We carry out the following 3 models:

- Model 1 only focuses on Panels B and C. This model only considers real data, i.e. banking data, to analyze stock market participation. Model 1 focuses on 77,365 retail clients.
- Model 2 focuses on Panels A (only "Risk tolerance"), B and C. Complementing Model 1, Model 2 considers retail clients' risk tolerance level in order to assess stock market participation. Due to unreported answers on risk tolerance, Model 2 focuses on 71,461 retail clients.
- Model 3 focuses on Panels A (only "Attitudes twd losses"), B and C. This model implements retail clients' decision-making during a hypothetical downturn in order to explain stock market participation. Model 3 focuses on 71,745 retail clients due to missing answers on attitudes towards losses.

We check the presence of possible multicollinearity problem by using the condition index of Belsley et al. (1980) and the Variance Inflation Factor (VIF). Respecting the critical threshold of both methods, we conclude we do not face such problem[24]. Average Marginal Effects (AMEs) are used for analyzing the magnitude effects.

Table 3 presents our findings. We first notice that almost all independent variables are statistically significant at reasonable significance levels and that implementing MiFID indicators (Models 2 and 3) into a basic model (Model 1) increases the quality of goodness-of-fit (see Appendix C for further explanation).

[24]Looking at the three models, the largest condition index is 17.76 which is below the critical threshold of 30. Besides, it also satisfies the threshold of 20 indicated by Erkel-Rousse (1995). As for the VIF, Chatterjee et al. (2000) denote that a VIF larger than 10 and/or a mean VIF greater than or equal to 2 indicate the presence of multicollinearity problem. In our empirical analysis, the largest VIF is 2.94 and the largest mean VIF is 1.54 (see Table 10 in Appendix B).

Table 3 – Stock market participation determinants

	Model 1		Model 2		Model 3	
	AMEs	**std**	**AMEs**	**std**	**AMEs**	**std**
Dependent variable						
Stocks						
Independent variables						
Panel A: MiFID indicators						
Risk tolerance						
0			(omitted)			
1			0.1000***	0.0022		
2			0.1821***	0.0053		
Attitudes twd losses						
1					-0.0817***	0.0049
2					-0.0215***	0.0048
3					(omitted)	
4					0.0633***	0.0037
Panel B: Socio-demographic indicators						
Gender	0.0146***	0.0021	0.0086***	0.0022	0.0127***	0.0023
Age	0.0037***	0.0001	0.0036***	0.0001	0.0038***	0.0001
Native	0.0454***	0.0033	0.0398***	0.0034	0.0444***	0.0035
Paris	0.0385***	0.0029	0.0368***	0.0030	0.0352***	0.0031
Matrimonial	0.0295***	0.0029	0.0224***	0.0030	0.0281***	0.0031
Self-employed	0.0091***	0.0031	0.0086***	0.0032	0.0096***	0.0033
Salaried	(omitted)		(omitted)		(omitted)	
Retired	-0.0215***	0.0033	-0.0189***	0.0034	-0.0216***	0.0035
No occupation	0.0118***	0.0039	0.0074*	0.0041	0.0119***	0.0042
Panel C: Wealth and patrimony indicators						
ln(Income)	0.0150***	0.0010	0.0087***	0.0010	0.0133***	0.0011
ln(Credit)	-0.0006***	0.0002	-0.0010***	0.0002	-0.0009***	0.0002
UL life insurance	0.1320***	0.0020	0.0985***	0.0022	0.1280***	0.0021
Retirement	0.0858***	0.0058	0.0737***	0.0059	0.0839***	0.0061
N	77,365		71,461		71,745	
LR Chi2	10,906.17		12,675.22		10,919.84	
Proba>chi2	0.0000		0.0000		0.0000	
Pseudo-R2	0.2028		0.2446		0.2102	
Log likelihood	-21,440.37		-19,576.17		-20,511.58	

Table 3 displays BLR results that aim to identify stock market participation determinants among retail clients of a large European retail bank. The dependent variable "Stocks" indicates whether a retail client directly or indirectly held at least one stock (1) or not (0) as of the 07/31/2015. Average marginal effects (AMEs) of independent variables are reported. Statistical significance levels are fixed at 1%, 5% and 10% and are denoted by ***, ** and * respectively.

Comparing Model 1 to Models 2 and 3, we point out that MiFID indicators exhibit the greatest magnitude effects compared to usual drivers of investment decisions, such as gender, age and income (Barber and Odean, 2001, Korniotis and Kumar, 2011). In addition, AMEs increase with the ordinal ranking of MiFID indicators. First, we find that high risk tolerant retail clients are more likely to invest in stocks. Likewise, Hong et al. (2004) show that stock ownership is greater among risk tolerant households in the US. Hoffmann et al. (2015) find that Dutch retail investors with higher levels of and upward revisions in risk tolerance are more likely to trade. Looking at attitudes towards losses, we find that retail clients willing to hold further financial securities during a downturn are more likely to hold stocks. Indeed, relative to the neutral position, AME is strongly negative for retail clients preferring to sell the entire portfolio (-8.17%) than those preferring to sell a part of the portfolio (-2.15%) and is strongly positive for those preferring to invest in further financial securities (6.33%). In other words, retail clients being less sensitive towards losses during the downturn have a high tendency to own stocks. In a similar vein, Dimmock and Kouwenberg (2010) show that the probability of investing in equity markets is lower among US households with high loss aversion. Finally, comparing the magnitude effects of the two attitudes, allows us to conclude that risk tolerance has a greater impact on the decision to hold stocks that the sensitivity towards losses.

In Panel B, across all models, we find that men are more likely to participate in stock market than women, like in Van Rooij et al. (2011) and Almenberg and Dreber (2015). Looking at AMEs, we find that being a man increases, on average, by 1.20% the likelihood to own stocks. Women have then a more conservative approach than men (Jacobsen et al., 2014). Older retail clients are more likely to invest in stocks than younger ones. This result is in line with Shum and Faig (2006). As for our specific variables, i.e. "Native", "Paris" and "Matrimonial", we point out that they all display greater AMEs than gender, age and professional occupations. We find that native-born retail clients are more likely to hold stocks and show for the first time that the country of birth has a strong impact on stockholding likelihood, like in the US. Specifically, US immigrants hold less financial assets, such as stocks and mutual funds compared to natives (Osili and Paulson, 2004, Chatterjee, 2009 and Luik and Steinhardt, 2016). As expected, geographic proximity with the capital region ("Paris") enhances stock market participation, a result in line with Arrondel et al. (2010). For the first time, we shed light on the contribution of matrimonial regime choice in understanding stockholding. Retail clients opting for the separation regime have a high tendency to hold stocks. As there is no joint-ownership between spouses, any gain and loss generated by stockholding does not impact the wealth of the other spouse. That financial independence promotes stock market participation.

Regarding professional occupations[25], we find that self-employed retail clients are more likely to own stocks than salaried employees. This result is consistent with the separation regime choice regarding financial independence. For salaried employees, financial constraint might limit stock market participation. Being retired

[25] In all models, "Salaried" is the reference category because it concentrates the highest number of retail clients. Besides, it is highly correlated with each of the three other occupations.

decreases, on average, by 2.07% the likelihood to own stocks compared to salaried employees. Being professionally active on labor market thus increases the tendency to participate in stock market as professional activity promotes social interactions. Recently, Fagereng et al. (2017) show that Norwegian households tend to reduce their portfolio share and exit the stock market around retirement. According to Hong et al. (2004), social households are more likely to invest in stock market than non-social households. They argue that individuals are attracted by stock market participation when many of their peers already invest. As for "No occupation", we find that retail clients exercising no professional activity are more likely to invest in stocks than salaried ones. This result is opposed to that of Grinblatt et al. (2011) who find that unemployed individuals display a lower stock market participation rate than employed ones. For these non-occupied individuals, we cannot make inferences on their social interactions and the only remaining explanation is the time free that they can allocate to stock investment.

In Panel C, we notice that the likelihood to invest in stock market increases, on average, by 0.12% given a 10% increase in the net monthly income. Our result is consistent with that of Barber and Odean (2001), Arrondel et al. (2010) and Liang and Guo (2015). Indeed, retail clients perceiving a high income level can afford to own stocks and to face some potential losses. As for credit, we find that the likelihood to invest in stock market decreases, on average, by 0.01% given a 10% increase in the credit amount remaining to reimburse. Indeed, being credit-constrained limits financial investment opportunities. As for risky financial products, we find that holding unit-linked life insurance and/or retirement plans affects positively the decision to hold stocks. Specifically, the magnitude effects of both products are higher than those of income whatever the model we focus on. The likelihood to hold stocks is, on average, 12% and 8% greater among retail clients holding unit-linked life insurance and retirement plans, respectively. These retail clients are then willing to accept a risk of loss in capital as both products depend on financial markets' performance. Further, they are more financially sophisticated due to their diversified portfolios (Boolell-Gunesh et al., 2009).

4 Robustness checks

In Section 4, we test the consistency of our findings by performing three robustness checks, noted RC1, RC2 and RC3 hereafter.

RC1 tests the impact of financial experience on the decision to own stocks. In Panel C, we introduce the variable "Account tenure" in order to measure the length of time (in years) during which an individual has been a retail client of the bank. It corresponds to the duration between the date of arrival of the client in the bank and the date of extraction of banking records, i.e. 07/31/2015. We then consider "Account tenure" as a proxy of retail clients' financial experience[26]. On average, individuals have been clients of the bank for 14 years. This period is much longer than those reported by Dorn and Sengmueller (2009) and Hoffmann et al.

[26]By using retail investors' answers, Glaser and Weber (2007) and Merkle (2015) introduce the variable "Experience" for referring the length of time during which retail investors have been directly investing in stock market.

(2015), i.e. 3 and 4 years, respectively. In RC1, “Age” is excluded due to its high correlation with “Account tenure” (Pearson correlation coefficient is 0.59).

RC2 tests whether stock market participation is impacted by retail client’s familiarity with the MiFID questionnaire content, by focusing on retail clients having answered thrice the questionnaire (N=11,839). We expect that a higher familiarity with the questionnaire indicates a higher confidence in the questionnaire objective, i.e. investor protection and suitability of advice, and a greater familiarity with stock markets. As expected, the stock holding rate of this subsample is about 30%, i.e. three times higher than that of the whole sample. We then check whether our findings are valid on this specific subsample.

RC3 tests the impact of financial literacy on stock market participation. According to the literature, financial literacy increases the likelihood to invest in stock (Van Rooij et al., 2011 and Mouna and Anis, 2017) and derivatives markets (Hsiao and Tsai, 2018). In Panel A, we introduce four variables extracted from the MiFID questionnaire answers. These “financial literacy variables” are coded 1 if retail clients declare they know the risk associated with stocks (“R_Stocks”), bonds (“R_Bonds”) or other financial products (“R_other” for unusual products such as warrants, deferred service settlements, convertible bonds or other financial investments), respectively. The last variable, “R_Markets”, is coded 1 if retail clients declare they understand the operation of financial markets. In our sample, about 61% of retail clients know the risk associated with stocks (N=71,880; std=0.49). 44% of retail clients know the risk associated with bonds (N=71,906; std=0.50). 20% of retail clients understand the functioning of financial markets (N=75,694; std=0.40). Only 17% of them know the risk associated with unusual financial products (N=50,411; std=0.37). We create four additional models because the four financial literacy variables are highly correlated and exclude “Risk tolerance” and “Attitudes twd losses” because of correlation problems with the financial literacy variables.

Robustness checks results are displayed in Table 4 (RC1 and RC2) and Table 5 (RC3).

Table 4 – RC1 and RC2 results

	RC1						RC2					
	Model 1		Model 2		Model 3		Model 1		Model 2		Model 3	
	AMEs	**std**	**AMEs**	**std**	**AMEs**	**std**	**AMEs**	**std**	**AMEs**	**std**	**AMEs**	**std**
Panel A: MiFID indicators												
Risk tolerance												
0			(omitted)						(omitted)			
1			0.0988***	0.0022					0.2033***	0.0077		
2			0.1817***	0.0052					0.3376***	0.0173		
Attitudes twd losses												
1					-0.0816***	0.0048					-0.1832***	0.0202
2					-0.0217***	0.0048					-0.0180	0.0172
3					(omitted)						(omitted)	
4					0.0603***	0.0037					0.1090***	0.0126
Panel B: Socio-demographic indicators												
Gender	0.0107***	0.0021	0.0050**	0.0022	0.0088***	0.0022	0.0176**	0.0080	0.0051	0.0077	0.0127	0.0080
Age							0.0065***	0.0003	0.0058***	0.0003	0.0064***	0.0003
Native	0.0290***	0.0033	0.0241***	0.0034	0.0276***	0.0035	0.0844***	0.0123	0.0727***	0.0118	0.0773***	0.0122
Paris	0.0345***	0.0029	0.0334***	0.0030	0.0315***	0.0031	0.0669***	0.0119	0.0687***	0.0115	0.0632***	0.0118
Matrimonial	0.0374***	0.0029	0.0300***	0.0030	0.0363***	0.0031	0.0596***	0.0113	0.0435***	0.0108	0.0543***	0.0112
Self-employed	0.0278***	0.0031	0.0266***	0.0032	0.0288***	0.0032	0.0286**	0.0120	0.0194*	0.0115	0.0244**	0.0119
Retired	0.0211***	0.0027	0.0227***	0.0028	0.0232***	0.0029	-0.0315***	0.0121	-0.0283**	0.0117	-0.0345***	0.0120
No occupation	0.0107***	0.0038	0.0070*	0.0040	0.0109***	0.0040	0.0165	0.0154	0.0041	0.0149	0.0145	0.0153
Salaried	(omitted)		(omitted)		(omitted)		(omitted)		(omitted)		(omitted)	
Panel C: Wealth and patrimony indicators												
ln(Income)	0.0198***	0.0010	0.0133***	0.0010	0.0183***	0.0011	0.0360***	0.0042	0.0231***	0.0037	0.0320***	0.0040
ln(Credit)	-0.0006***	0.0002	-0.0009***	0.0002	-0.0008***	0.0002	-0.0014*	0.0008	-0.0018**	0.0008	-0.0017**	0.0008
Account tenure	0.0039***	0.0001	0.0039***	0.0001	0.0040***	0.0001						
UL life insurance	0.1186***	0.0020	0.0861***	0.0022	0.1150***	0.0021	0.2039***	0.0069	0.1360***	0.0073	0.1901***	0.0070
Retirement	0.0777***	0.0057	0.0655***	0.0059	0.0757***	0.0060	0.1356***	0.0174	0.1089***	0.0167	0.1260***	0.0173
N	77,365		71,461		71,745		11,839		11,818		11,824	
LR Chi2	11,762.76		13,481.78		11,718.33		2,198.01		2,960.72		2,377.72	
Prob>Chi2	0.0000		0.0000		0.0000		0.0000		0.0000		0.0000	
Pseudo-R2	0.2187		0.2601		0.2256		0.1526		0.2058		0.1652	
Log likelihood	-21,012.08		-19,172.89		-20,112.34		-6,103.38		-5,712.89		-6,007.37	

Table 4 reports robustness checks 1 and 2 results. Robustness check 1 (RC1) considers “Account tenure” instead of “Age” and focuses on the whole sample of retail clients. Robustness check 2 (RC2) refers to a subsample of retail clients having answered thrice the MiFID questionnaire and considers the same independent variables included in our empirical analyses. Statistical significance levels are fixed at 1%, 5% and 10% and are denoted by ***, ** and * respectively.

Table 5 – RC3 results

	RC3							
	Model 4		Model 5		Model 6		Model 7	
	AMEs	std	AMEs	std	AMEs	std	AMEs	std
				Panel A: MiFID indicators				
R_Stocks	0.1227***	0.0035						
R_Bonds			0.0893***	0.0024				
R_Markets					0.0793***	0.0021		
R_other							0.0868***	0.0034
				Panel B: Socio-demographic indicators				
Gender	0.0142***	0.0022	0.0122***	0.0022	0.0119***	0.0021	0.0119***	0.0031
Age	0.0034***	0.0001	0.0034***	0.0001	0.0035***	0.0001	0.0044***	0.0001
Native	0.0350***	0.0035	0.0376***	0.0035	0.0404***	0.0033	0.0456***	0.0048
Paris	0.0334***	0.0031	0.0311***	0.0031	0.0344***	0.0029	0.0364***	0.0042
Matrimonial	0.0254***	0.0030	0.0232***	0.0030	0.0247***	0.0029	0.0284***	0.0041
Self-employed	0.0061*	0.0032	0.0073**	0.0032	0.0063**	0.0031	0.0087**	0.0044
Salaried	(omitted)		(omitted)		(omitted)		(omitted)	
Retired	-0.0190***	0.0035	-0.0199***	0.0035	-0.0213***	0.0033	-0.0211***	0.0047
No occupation	0.0149***	0.0042	0.0108***	0.0041	0.0093**	0.0039	0.0143**	0.0057
				Panel C: Wealth and patrimony indicators				
ln(Income)	0.0091***	0.0010	0.0091***	0.0010	0.0104***	0.0010	0.0097***	0.0013
ln(Credit)	-0.0010***	0.0002	-0.0009***	0.0002	-0.0007***	0.0002	-0.0017***	0.0003
UL life insurance	0.1186***	0.0021	0.1225***	0.0021	0.1188***	0.0020	0.1555***	0.0028
Retirement	0.0800***	0.0060	0.0779***	0.0060	0.0775***	0.0058	0.1035***	0.0080
N	71,880		71,906		75,694		50,411	
LR Chi2	11,817.60		11,710.49		12,043.25		8,048.58	
Prob>Chi2	0.0000		0.0000		0.0000		0.0000	
Pseudo-R2	0.2273		0.2252		0.2261		0.1834	
Log likelihood	-20,087.64		-20,148.47		-20,609.89		-17,921.81	

Table 5 reports RC3 results. Belonging to Panel A, four financial literacy variables, i.e. “R_Stocks”, “R_Bonds”, “R_Markets” and “R_other”, are separately treated in RC3 and are introduced in Model 4, Model 5, Model 6 and Model 7 respectively. Statistical significance levels are fixed at 1%, 5% and 10% and are denoted by ***, ** and * respectively.

In RC1, we find that retail clients with high account tenure, i.e. financial experience, are more likely to invest in stocks. Furthermore, "Account tenure" is significant at all reasonable levels and displays AMEs close to those of "Age" (Table 3). As a result, we conclude that "Age" and "Account tenure" are two equivalent experience proxies. In a similar vein, Hoffmann et al. (2015) find that older and more experienced investors are more likely to trade derivatives. Our results are different from the ones of Korniotis and Kumar (2011) who showed that introducing both age and experience allows distinguishing two confounding effects where the negative effects of age dominate the positive effects of experience. For these authors, age refers to cognitive aging, i.e. the weakening of memory with age, whereas experience refers to accumulation of greater investment knowledge with age. These interesting effects are not detected in our data.

In RC2, AMEs of quite all independent variables are greater than those corresponding to the whole sample. However, "Attitudes twd losses" coded 2, "Gender" and "No occupation" (Models 2 and 3) are not significant. As RC2 reinforces our findings, we conclude that administering several times the MiFID questionnaire is useful for threefold reasons. First, in the third questionnaire, retail clients are more familiar and confident with the MiFID questionnaire. Confidence may explain to a great extend the decrease of the number of unreported answers between two successive questionnaires[27]. Second, investment service providers are more likely to offer suited advice and financial products to their clients as they can collect further information on them. Finally, administering several times the MiFID questionnaire may also improve retail clients' financial trading activity[28].

In RC3, we notice that financial literacy variables are significant at all reasonable levels. The likelihood to hold stocks is 12.27%, 8.93% and 8.68% higher among retail clients who know the risk associated with stocks, bonds and other financial products, respectively. It is 7.93% higher among retail clients who understand the functioning of financial markets. Financial literacy has then a positive impact on stock market participation. Moreover, all independent variables of Panels B and C have identical impact whatever the financial literacy variable introduced. As the impact of age is the same in the four models relative to our main findings, we conclude that age, i.e. experience, completes financial literacy. Therefore, RC3 helps to distinguish the positive impacts of financial literacy and experience in explaining stock market participation. Finally, implementing financial literacy variables allows to increase pseudo-R^2 value, except for "R_other" (Model 7). As financial literacy variables are extracted from the MiFID questionnaire answers, we anew shed light on the contribution of the use of MiFID variables in explaining stock market participation.

[27] There are, on average, 8.29% of unreported answers in the first questionnaire, 1.48% in the second questionnaire and 0.15% in the third questionnaire.

[28] For example, a retail client who does not know what a stock is in the first questionnaire may try to get information on it.

5 Conclusion

The Markets in Financial Instrument Directive (MiFID) requires investment service providers to get a thorough knowledge of their clients by administering a questionnaire, called *MiFID questionnaire*, to offer advice and financial products perfectly suited to their situation. In this paper, we analyze stock market participation of more than 70,000 retail clients of a large European retail bank by using two datasets, MiFID questionnaire answers and banking records.

We find that the likelihood to invest in stocks increases with risk tolerance and with a low sensitivity towards losses. These self-assessed attitudes display greater magnitude effects than usual drivers of investment decisions, such as gender, age and income. In addition, we show that the country of birth and residency, and matrimonial regime choice have a strong impact on the decision to hold stocks. The likelihood to invest in stocks is higher among natives and those living in the biggest region of the country. The matrimonial separation regime demonstrates the willing to be financially independent which promotes stock market participation. Regarding wealth and patrimony indicators, we find that holding other risky financial products increases the propensity to own stocks more than income does. Finally, we show that our findings are still valid on a specific subsample composed of retail clients who are familiar both with the MiFID questionnaire content and with the stock market. We also show that highly literate individuals are more likely to participate on the stock market than low literate ones.

Our results contribute to the actual debate between professionals, regulators and academics about the usefulness of MiFID indicators. de Palma and Picard (2010) have realized a first diagnosis of 14 MiFID questionnaires provided by 10 financial intermediaries in France. They suggested that a quantitative measurement of risk-taking preferences is necessary for ensuring that investment service providers offer suited advice and financial products to their clients. This claim is now achieved as preferences assessed in the MiFID questionnaire explain stock market participation.

Further, this paper helps giving insights to improve MiFID questionnaire. As there is no regulatory constraint on the MiFID questionnaire content, investment service providers are free to establish their questionnaire. Therefore, questionnaire length may be different from a bank to another one. When they exist, questions dealing with individual characteristics (e.g. gender, age and marital status) are, in general, located at the beginning of the questionnaire whereas more important information (risk tolerance and attitudes towards losses) are in the end. However, these pieces of information are already recorded in the bank database. Even if each question brings additional information about retail clients, fatigue effect or mistrust should be considered and the likelihood of getting high data reliability decreases with questionnaire length. Besides, questionnaire length may also influence the number of unreported answers which may restrict the ability of investment service providers to offer suited advice and financial products to their clients. For these reasons, we suggest reducing the length of the questionnaire. We also recommend that MiFID questionnaire mainly focus on retail clients' own preferences which have been demonstrated as key drivers of stock investment decision in this study.

References

Ackert, L., Church, B. K. and Englis, B. (2002). The asset allocation decision and investor heterogeneity: a puzzle? *Journal of Economic Behavior and Organization* 47: 423–433.

Agnew, J., Balduzzi, P. and Sunden, A. (2003). Portfolio choice and trading in a large 401 (k) plan. *The American Economic Review* 93: 193–215.

Aitken, M., Cumming, D. and Zhan, F. (2015). Exchange trading rules, surveillance and suspected insider trading. *Journal of Corporate Finance* 34: 311 – 330.

Almenberg, J. and Dreber, A. (2015). Gender, stock market participation and financial literacy. *Economics Letters* 137: 140 – 142.

Antoniou, C., Harris, R. D. and Zhang, R. (2015). Ambiguity aversion and stock market participation: An empirical analysis. *Journal of Banking & Finance* 58: 57 – 70.

Aouadi, A., Arouri, M. and Teulon, F. (2013). Investor attention and stock market activity: Evidence from france. *Economic Modelling* 35: 674 – 681.

Arrondel, L., Debbich, M. and Savignac, F. (2015). Stockholding in France: the role of financial literacy and information. *Applied Economics Letters* 22: 1315–1319.

Arrondel, L., Pardo, H. C. and Oliver, X. (2010). Temperance in stock market participation: Evidence from france. *Economica* 77: 314–333.

Attanasio, O., Banks, J. and Tanner, S. (2002). Asset Holding and Consumption Volatility. *Journal of Political Economy* 110: 771–792.

Bajtelsmit, V. L., Bernasek, A. and Jianakoplos, N. A. (1999). Gender differences in defined contribution pension decisions. *Financial Services Review* 8: 1–10.

Bajtelsmit, V. L. and VanDerhei, J. L. (1997). Risk aversion and pension investment choices. *Positioning pensions for the twenty-first century* 45: 66.

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

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

Bakshi, G. S. and Chen, Z. (1994). Baby boom, population aging, and capital markets. *Journal of Business* 67: 163–202.

Balloch, A., Nicolae, A. and Philip, D. (2014). Stock market literacy, trust, and participation. *Review of Finance* 19: 1925–1963.

Barber, B. M. and Odean, T. (2001). Boys will be boys: Gender, overconfidence, and common stock investment. *Quarterly journal of Economics* : 261–292.

Bauer, R., Cosemans, M. and Eichholtz, P. (2009). Option trading and individual investor performance. *Journal of Banking and Finance* 33: 731–746.

Becker, T. and Shabani, R. (2010). Outstanding debt and the household portfolio. *Review of Financial Studies* 23: 2900–2934.

Bellofatto, A., De Winne, R. and D'Hondt, C. (2014). Beyond the Disposition Effect: Evidence from the 1999-2012 period. In *The 2014 EFM Merton H Miller Doctoral Seminar of the European Financial Management Association (EFMA)*.

Bellofatto, A., De Winne, R. and D'Hondt, C. (2015). To what extent can MiFID tests be informative? In *Southern Finance Association 2015 Annual Meeting*.

Belsley, D. A., Kuh, E. and Welsch, R. E. (1980). Regression diagnostics: Identifying influential observations and sources of collinearity. New York: John Wiley.

Bernasek, A. and Shwiff, S. (2001). Gender, Risk, and Retirement. *Journal of Economic Issues* 35: 345–356.

Bertocchi, G., Brunetti, M. and Torricelli, C. (2011). Marriage and other risky assets: A portfolio approach. *Journal of Banking and Finance* 35: 2902–2915.

Bianchi, M. (2018). Financial literacy and portfolio dynamics. *The Journal of Finance* forthcoming.

Blau, B. M. (2017). Skewness preferences, asset prices and investor sentiment. *Applied Economics* 49: 812 – 822.

Bodie, Z. and Crane, D. B. (1997). Personal Investing: Advice, Theory, and Evidence from a Survey of TIAA-CREF Participants. SSRN Scholarly Paper ID 36158, Social Science Research Network, Rochester, NY.

Bogan, V. (2008). Stock market participation and the internet. *Journal of Financial & Quantitative Analysis* 43: 191 – 211.

Bonaparte, Y. and Kumar, A. (2013). Political activism, information costs, and stock market participation. *Journal of Financial Economics* 107: 760 – 786.

Booij, A. S. and Kuilen, G. Van de (2009). A parameter-free analysis of the utility of money for the general population under prospect theory. *Journal of Economic psychology* 30: 651–666.

Boolell-Gunesh, S., Broihanne, M.-H. and Merli, M. (2009). Disposition effect, investor sophistication and taxes: Some French Specificities. *Finance* 30: 51–78.

Boolell-Gunesh, S., Broihanne, M.-H. and Merli, M. (2012). Sophistication of Individual Investors and Disposition Effect Dynamics. *Finance* 33: 9–37.

Booth, A. and Nolen, P. (2012). Salience, risky choices and gender. *Economics Letters* 117: 517–520.

Bricker, J., Dettling, L. J., Henriques, A., Hsu, J. W., Moore, K. B., Sabelhaus, J., Thompson, J. and Windle, R. A. (2014). Changes in us family finances from 2010 to 2013: Evidence from the survey of consumer finances. *Federal Reserve Bulletin* 100: 1–41.

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

Brooks, C., Sangiorgi, I., Hillenbrand, C. and Money, K. (2018). Why are older investors less willing to take financial risks? *International Review of Financial Analysis* 56: 52 – 72.

Brown, G. W. and Cliff, M. T. (2005). Investor sentiment and asset valuation. *The Journal of Business* 78: 405–440.

Cahuzac, E. and Bontemps, C. (2008). *Stata par la pratique : statistiques, graphiques et elements de programmation*. No. splp in Stata Press books. StataCorp LP.

Cardak, B. A. and Wilkins, R. (2009). The determinants of household risky asset holdings: Australian evidence on background risk and other factors. *Journal of Banking and Finance* 33: 850 – 860.

Chang, C.-C., Hsieh, P.-F. and Wang, Y.-H. (2015). Sophistication, sentiment, and misreaction. *Journal of Financial & Quantitative Analysis* 50: 903 – 928.

Chang, T. Y., Solomon, D. H. and Westerfield, M. M. (2016). Looking for someone to blame: Delegation, cognitive dissonance, and the disposition effect. *The Journal of Finance* 71: 267–302.

Charness, G. and Gneezy, U. (2012). Strong evidence for gender differences in risk taking. *Journal of Economic Behavior & Organization* 83: 50 – 58, gender Differences in Risk Aversion and Competition.

Chatterjee, S. (2009). Individual Stockownership in the United States: Native-Immigrant Gap and the Role of Risk Tolerance. *International Research Journal of Finance and Economics* 28: 160–168.

Chatterjee, S., Hadi, A. S. and Price, B. (2000). *Regression Analysis by Example*. New York : Chapman & Hall.

Chen, G., Kim, K. A., Nofsinger, J. R. and Rui, O. M. (2007). Trading performance, disposition effect, overconfidence, representativeness bias, and experience of emerging market investors. *Journal of Behavioral Decision Making* 20: 425–451.

Chen, G.-M., Kim, K. A., Nofsinger, J. R. and Rui, O. M. (2004). Behavior and performance of emerging market investors: Evidence from China. *unpublished Washington State University Working paper (January)* .

Cho, I. (2014). Homeownership and investment in risky assets in europe. *Review of European Studies* 6.

Christelis, D., Jappelli, T. and Padula, M. (2010). Cognitive abilities and portfolio choice. *European Economic Review* 54: 18 – 38.

Constantinides, G., Donaldson, J. B. and Mehra, R. (2002). Junior can't borrow: A new perspective on the equity premium puzzle. *The Quarterly Journal of Economics* 117: 269–296.

Cronqvist, H., Previtero, A., Siegel, S. and White, R. E. (2016). The fetal origins hypothesis in finance: Prenatal environment, the gender gap, and investor behavior. *Review of Financial Studies* 29: 739 – 786.

Cumming, D., Johan, S. and Li, D. (2011). Exchange trading rules and stock market liquidity. *Journal of Financial Economics* 99: 651 – 671.

Cummings, R. G., Brookshire, D. S., Bishop, R. C. and Arrow, K. J. (1986). Valuing environmental goods: an assessment of the contingent valuation method.

Daniel, K. and Hirshleifer, D. (2015). Overconfident investors, predictable returns, and excessive trading. *Journal of Economic Perspectives* 29: 61–88.

De Bourmont, M. (2012). La résolution d'un problème de multicolinéarité au sein des études portant sur les déterminants d'une publication volontaire d'informations : proposition d'un algorithme de décision simplifié basé sur les indicateurs de Belsley, Kuh et Welsch (1980). In *Comptabilités et innovation*. Grenoble, France, cd–rom.

Dhar, R. and Zhu, N. (2006). Up Close and Personal: Investor Sophistication and the Disposition Effect. *Management Science* 52: 726–740.

D'Hondt, C. and Roger, P. (2018). Investor sentiment and return predictability: The power of ignorance. *Finance* 38: forthcoming.

Dimmock, S. G. and Kouwenberg, R. (2010). Loss-aversion and household portfolio choice. *Journal of Empirical Finance* 17: 441–459.

Dobni, D. M. and Racine, M. D. (2015). Stock market image: The good, the bad, and the ugly. *Journal of Behavioral Finance* 16: 130–139.

Dobni, D. M. and Racine, M. D. (2016). Investors' images of the stock market: Antecedents and consequences. *Financial Services Review* 25: 1 – 28.

Dorn, D. and Huberman, G. (2005). Talk and Action: What Individual Investors Say and What They Do. *Review of Finance* 9: 437–481.

Dorn, D. and Sengmueller, P. (2009). Trading as entertainment? *Management Science* 55: 591–603.

Dwyer, P. D., Gilkeson, J. H. and List, J. (2002). Gender differences in revealed risk taking: evidence from mutual fund investors. *Economics Letters* 76: 151–158.

Erkel-Rousse, H. (1995). Detection de la multicolinearite dans un modele lineaire ordinaire: quelques elements pour un usage averti des indicateurs de belsley, kuh et welsch. *Revue de Statistique Appliquee* 43: 19–42.

Faff, R., Mulino, D. and Chai, D. (2008). On the linkage between financial risk tolerance and risk aversion. *Journal of Financial Research* 31: 1–23.

Fagereng, A., Gottlieb, C. and Guiso, L. (2017). Asset market participation and portfolio choice over the life-cycle. *The Journal of Finance* 72: 705–750.

Fan, J. X. and Xiao, J. J. (2006). Cross-cultural differences in risk tolerance: A comparison between Chinese and Americans. *Journal of Personal Finance* 5: 54–75.

Feng, L. and Seasholes, M. S. (2005). Do investor sophistication and trading experience eliminate behavioral biases in financial markets? *Review of Finance* 9: 305–351.

Feng, L. and Seasholes, M. S. (2008). Individual investors and gender similarities in an emerging stock market. *Pacific-Basin Finance Journal* 16: 44–60.

Fischer, M. and Jensen, B. A. (2015). Taxation, transfer income and stock market participation. *Review of Finance* 19: 823 – 863.

Foucault, T., Sraer, D. and Thesmar, D. (2011). Individual investors and volatility. *Journal of Finance* 66: 1369–1406.

Fratantoni, M. C. (1998). Homeownership and investment in risky assets. *Journal of Urban Economics* 44: 27–42.

Fuertes, A.-M., Muradoglu, G. and Ozturkkal, B. (2014). A behavioral analysis of investor diversification. *The European Journal of Finance* 20: 499–523.

Georgarakos, D. and Inderst, R. (2014). Financial Advice and Stock Market Participation. CEPR Discussion Papers 9922, C.E.P.R. Discussion Papers.

Glaser, M. and Weber, M. (2007). Overconfidence and trading volume. *The Geneva Risk and Insurance Review* 32: 1–36.

Goetzmann, W. and Kumar, A. (2008). Equity Portfolio Diversification. *Review of Finance* 12: 433–463.

Grable, J. E. (2000). Financial risk tolerance and additional factors that affect risk taking in everyday money matters. *Journal of Business and Psychology* 14: 625–630.

Grable, J. E. and Joo, S.-H. (2004). Environmental and biophysical factors associated with financial risk tolerance. *Journal of Financial Counseling and Planning* 15.

Grinblatt, M. and Keloharju, M. (2009). Sensation Seeking, Overconfidence, and Trading Activity. *Journal of Finance* 64: 549–578.

Grinblatt, M., Keloharju, M. and Linnainmaa, J. (2011). Iq and stock market participation. *The Journal of Finance* 66: 2121–2164.

Guiso, L. and Jappelli, T. (2005). Awareness and stock market participation. *Review of Finance* 9: 537–567.

Guiso, L., Jappelli, T. and Terlizzese, D. (1996). Income Risk, Borrowing Constraints, and Portfolio Choice. *American Economic Review* 86: 158–172.

Guiso, L., Sapienza, P. and Zingales, L. (2008). Trusting the stock market. *The Journal of Finance* 63: 2557–2600.

Guiso, L. and Sodini, P. (2013). *Household Finance: An Emerging Field*. Elsevier, Handbook of the Economics of Finance *2*, chap. 0. 1397–1532.

Halek, M. and Eisenhauer, J. G. (2001). Demography of risk aversion. *The Journal of Risk and Insurance* 68: 1–24.

Haliassos, M. and Bertaut, C. C. (1995). Why do so few hold stocks? *Economic Journal* 105: 1110–29.

Halko, M.-L., Kaustia, M. and Alanko, E. (2012). The gender effect in risky asset holdings. *Journal of Economic Behavior & Organization* 83: 66 – 81.

Hallahan, T. A., Faff, R. W. and McKenzie, M. D. (2004). An empirical investigation of personal financial risk tolerance. *Financial Services Review* : 57–78.

Heaton, J. and Lucas, D. (2000). Portfolio choice and asset prices: The importance of entrepreneurial risk. *Journal of Finance* 55: 1163–1198.

Hinz, R. P., McCarthy, D. D. and Turner, J. A. (1997). Are Women Conservative Investors? Gender Differences in Participant Directed Pension Investments. Pension Research Council Working Papers 96-17, Wharton School Pension Research Council, University of Pennsylvania.

Hirshleifer, D. (2001). Investor psychology and asset pricing. *The Journal of Finance* 56: 1533–1597.

Hoffmann, A. O., Post, T. and Pennings, J. M. (2013). Individual investor perceptions and behavior during the financial crisis. *Journal of Banking and Finance* 37: 60–74.

Hoffmann, A. O., Post, T. and Pennings, J. M. (2015). How investor perceptions drive actual trading and risk-taking behavior. *Journal of Behavioral Finance* 16: 94–103.

Holt, C. and Laury, S. (2002). Risk aversion and incentive effects. *American Economic Review* 92: 1644–1655.

Hong, H., Kubik, J. D. and Stein, J. C. (2004). Social Interaction and Stock-Market Participation. *The Journal of Finance* 59: 137–163.

Hsiao, Y.-J. and Tsai, W.-C. (2018). Financial literacy and participation in the derivatives markets. *Journal of Banking & Finance* 88: 15 – 29.

Jacobsen, B., Lee, J. B., Marquering, W. and Zhang, C. Y. (2014). Gender differences in optimism and asset allocation. *Journal of Economic Behavior and Organization* 107: 630–651.

Jianakoplos, N. and Bernasek, A. (1998). Are Women More Risk Averse? *Economic Inquiry* 36: 620–30.

Kaustia, M. and Torstila, S. (2011). Stock market aversion? political preferences and stock market participation. *Journal of Financial Economics* 100: 98–112.

Korniotis, G. M. and Kumar, A. (2011). Do Older Investors Make Better Investment Decisions? *The Review of Economics and Statistics* 93: 244–265.

Kremelberg, D. (2011). *Practical Statistics: A Quick and Easy Guide to IBM SPSS Statistics, STATA, and Other Statistical Software*. SAGE Publications.

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

Kumar, A. (2009). Who gambles in the stock market? *The Journal of Finance* 64: 1889–1933.

Lease, R. C., Lewellen, W. G. and Schlarbaum, G. G. (1974). The Individual Investor: Attributes and Attitudes. *The Journal of Finance* 29: 413–433.

Liang, P. and Guo, S. (2015). Social interaction, internet access and stock market participation-an empirical study in china. *Journal of Comparative Economics* 43: 883 – 901.

Lintner, J. (1965). The valuation of risk assets and the selection of risky investments in stock portfolios and capital budgets. *The Review of Economics and Statistics* 47: 13–37.

Long, J. S. and Freese, J. (2006). *Regression Models for Categorical Dependent Variables using Stata, 2nd Edition*. StataCorp LP, 2nd ed.

Luik, M.-A. and Steinhardt, M. F. (2016). Immigrant-native differences in stockholding - the role of cognitive and non-cognitive skills. *Journal of Empirical Finance* 38: 103 – 119.

Maccrimmon, K. and Wehrung, D. A. (1986). *The management of uncertainty: Taking risks.*. New york ed.

Magron, C. (2012). Performance of individual investors and personal investment objectives. Working Papers of LaRGE Research Center 2012-07, Laboratoire de Recherche en Gestion et Economie (LaRGE), Universite de Strasbourg.

Magron, C. (2014). The trading performance of individual investors. *Bankers, Markets and Investors* : 43–52.

Mankiw, N. G. and Zeldes, S. P. (1991). The consumption of stockholders and nonstockholders. *Journal of Financial Economics* 29: 97–112.

Mazzoli, C. and Marinelli, N. (2014). Determinants of risk-suitable investment portfolios: Evidence from a sample of italian householders. *Journal of Economic & Financial Studies* 2: 50–63.

Merkle, C. (2015). Financial loss aversion illusion. *Browser Download This Paper* .

Merli, M. and Roger, T. (2013). What drives the herding behavior of individual investors? *Finance* 34: 67–104.

Mitchell, R. C. and Carson, R. T. (1989). Using surveys to value public goods: the contingent valuation method.

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

Mossin, J. (1966). Equilibrium in a capital asset market. *Econometrica* 34: 768–783.

Mouna, A. and Anis, J. (2017). Financial literacy in tunisia: Its determinants and its implications on investment behavior. *Research in International Business and Finance* 39: 568 – 577.

Odean, T. (1999). Do Investors Trade Too Much? *American Economic Review* 89: 1279–1298.

Osili, U. O. and Paulson, A. (2004). Immigrant-native differences in financial market participation. *Federal Reserve Bank of Chicago, WP* 18.

Palma, A. de and Picard, N. (2010). Evaluation of MiFID qquestionnaire in France. Tech. rep., AMF, study for the Autorite des Marches Financiers.

Peress, J. (2005). Information vs. entry costs: What explains u.s. stock market evolution?. *Journal of Financial & Quantitative Analysis* 40: 563 – 594.

Poterba, J. M. and Samwick, A. A. (1995). Stock Ownership Patterns, Stock Market Fluctuations, and Consumption. *Brookings Papers on Economic Activity* 26: 295–372.

Powell, M. and Ansic, D. (1997). Gender differences in risk behaviour in financial decision-making: An experimental analysis. *Journal of Economic Psychology* 18: 605–628.

Richards, D. W., Rutterford, J., Kodwani, D. and Fenton-O'Creevy, M. (2017). Stock market investors' use of stop losses and the disposition effect. *The European Journal of Finance* 23: 130–152.

Riley, W. B. and Chow, K. V. (1992). Asset allocation and individual risk aversion. *Financial Analysts Journal* 48: 32–37.

Roger, P. (2014). The 99% Market Sentiment Index. *Finance* 35: 53–96.

Ryu, D., Kim, H. and Yang, H. (2017). Investor sentiment, trading behavior and stock returns. *Applied Economics Letters* 24: 826–830.

Sharpe, W. F. (1964). Capital asset prices: A theory of market equilibrium under conditions of risk. *The Journal of Finance* 19: 425–442.

Shum, P. and Faig, M. (2006). What explains household stock holdings? *Journal of Banking and Finance* 30: 2579 – 2597.

Siwar, E. (2011). Survey of the phenomenon of overreaction and underreaction on french stock market. *IUP Journal of Behavioral Finance* 8: 23 – 46.

Sung, J. and Hanna, S. (1996). Factors related to risk tolerance. *Journal of Financial Counseling and Planning* 7: 11.

Tekce, B. and Yilmaz, N. (2015). Are individual stock investors overconfident? evidence from an emerging market. *Journal of Behavioral and Experimental Finance* 5: 35 – 45.

Tversky, A. and Kahneman, D. (1992). Advances in prospect theory: Cumulative representation of uncertainty. *Journal of Risk and Uncertainty* 5: 297–323.

Van Rooij, M., Lusardi, A. and Alessie, R. (2011). Financial literacy and stock market participation. *Journal of Financial Economics* 101: 449 – 472.

Vissing-Jorgensen, A. (2002). Towards an explanation of household portfolio choice heterogeneity: Nonfinancial income and participation cost structures. Tech. rep., National Bureau of Economic Research.

Wachter, J. A. and Yogo, M. (2010). Why do household portfolio shares rise in wealth? *Review of Financial Studies* 23: 3929–3965.

Wang, H. and Hanna, S. D. (1997). Does risk tolerance decrease with age? *Financial Counseling and Planning* 8: 27–31.

Weber, M., Weber, E. U. and Nosic, A. (2013). Who takes risks when and why: Determinants of changes in investor risk taking. *Review of Finance* 17: 847–883.

Weber, M. and Welfens, F. (2007). An Individual Level Analysis of the Disposition Effect: Empirical and Experimental Evidence. Sonderforschungsbereich 504 Publications 07-45, Sonderforschungsbereich 504, Universitat Mannheim.

Yang, C. and Li, J. (2013). Investor sentiment, information and asset pricing model. *Economic Modelling* 35: 436 – 442.

Appendix

A Measures of association between variables

We study the relationships existing between variables as defined in Table 1. This preliminary step is necessary to build our estimation models carefully. We use 4 appropriate measures to analyze associations between variables: Phi coefficient (Φ), Cramer's V (V)[29], point biserial correlation coefficient (r_{pb}) and Spearman's rank correlation coefficient (r_{sp}).

We use Phi coefficient for studying the association between binary variables, Cramer's V for studying nominal variables (at least one of which has more than two modalities), point biserial correlation coefficient for binary/continuous variables and Spearman's rank correlation coefficient for binary/ordinal and continuous/ordinal variables. Contrary to the other association measures, Spearman's rank correlation is a non-parametric measure that does not assume normal distribution and linear relationship between variables. It is used for variables with ordinal measurement.

These association measures are presented separately for a total of 71,188 retail clients for which data are available. We pay attention to binary associations having significant coefficients exceeding 0.20 in absolute value and we carefully study the potential for multicollinearity issues between independent variables if coefficients exceed 0.50 (De Bourmont, 2012) in Section 3.

- *Phi coefficient*

Following Kremelberg (2011), we report the adjusted Phi value instead of Phi coefficient value due to unequal marginal distributions of binary variables[30]. Table 6 displays all binary pairwise variables.

Table 6 – Phi coefficients

	Gender	Native	Paris	Matrimonial	Stocks	UL life insurance	Retirement
Gender	1						
Native	0.05***	1					
Paris	0.03***	-0.21***	1				
Matrimonial	0.06***	0.01***	0.05***	1			
Stocks	0.05***	0.06***	0.05***	0.09***	1		
UL life insurance	0.02***	0.04***	0.03***	0.07***	0.41***	1	
Retirement	0.08**	0.03***	0.06***	0.06***	0.33***	0.50***	1

Table 6 presents Phi coefficients for all binary variables pairs. Chi-squared test is used for assessing coefficients significance. Statistical significance levels are fixed at 1%, 5% and 10% and are denoted by ***, ** and *, respectively.

[29] Unlike the three binary association measures, Cramer's V is comprised between 0 and 1.

[30] Phi coefficient value is equivalent to Pearson correlation coefficient for binary variables. It represents the square root of Chi-squared statistic divided by the sample size. However, it is assumed that each category approximately contains 50% of individuals. By violating this condition, Phi coefficient maximum value may be lower than 1 leading thus to misinterpretation of the strength of the binary association. For getting Phi adjusted value, we divided Phi value by its maximum possible value. This maximum value is computed by using the following formula : $\frac{\sqrt{p_j - p_j p_i}}{\sqrt{p_i - p_i p_j}}$ where $p_j$is the row or column containing the lowest proportion and p_i is the row or column containing the second lowest proportion. Thereby, this measure takes into account marginal distribution of each binary variable. Like Pearson correlation coefficient, the upper bound of this adjusted value is fixed to 1.

Table 6 shows that “Native” and “Paris” are negatively associated (Φ=-0.21). We also notice that there is a strong association between risky financial products. We can hypothesize that the probability of stockholding increases with the probability of holding other financial products. Indeed, our main variable “Stocks” is positively associated with “UL life insurance” (Φ=0.41) and to “Retirement” (Φ=0.33). In the same manner, “UL life insurance” and “Retirement” display a high Phi coefficient value (Φ=0.50). These relationships are consistent since all these financial products are risky investments.

- *Cramer’s V*

Cramer’s V allows to assess relationships between professional occupations (“Occupations”) and binary variables. Results are displayed in Table 7.

Table 7 – Cramer’s V coefficients

	Occupations
Gender	0.18***
Native	0.03***
Paris	0.04***
Matrimonial	0.17***
Stocks	0.17***
UL life insurance	0.14***
Retirement	0.05***

Table 7 presents Cramer’s V coefficients corresponding to each binary/nominal qualitative pairwise variables. Chi-squared test is used for assessing coefficients significance. Statistical significance levels are fixed at 1%, 5% and 10% and are denoted by ***, ** and *, respectively.

In Table 7, we notice that there is no strong association between professional occupations and binary variables. Indeed, the highest values are 0.18 and 0.17 indicating that professional occupations are weakly associated with “Gender”, “Matrimonial” and “Stocks”.

- *Point biserial correlation coefficient*

This coefficient is computed for assessing relationships between age and binary variables. Results are reported in Table 8.

Table 8 – Point biserial correlation coefficients

	Age
Gender	-0.05***
Native	-0.03***
Paris	0.00
Matrimonial	0.10***
Stocks	0.26***
UL life insurance	0.22***
Retirement	0.03***

Table 8 displays point biserial correlation coefficients corresponding to each binary/continuous pairwise variables. t-test is used for assessing coefficients significance. Statistical significance levels are fixed at 1%, 5% and 10% and are denoted by ***, ** and *, respectively.

Table 8 indicates that age is positively correlated to "Stocks" (r_{pb}=0.26) and "UL life insurance" (r_{pb} =0.22) leading us to hypothesize that stockholding increases with age.

- *Spearman rank correlation coefficient*

By applying this association measure, we look at the relationship between ordinal qualitative variables (i.e. MiFID indicators, "Income" and "Credit") and banking records, including our main variable "Stocks". We note that "Credit" and "Income" are ordinal qualitative variables. For these two variables, we use their numerical modalities (indicated in parentheses and superscript in Table 2) instead of their monetary codes (see "CODES" in Table 2).

Table 9 – Spearman rank correlation coefficients

	Risk tolerance	**Attitudes twd losses**	**Income**	**Credit**
Risk tolerance	1			
Attitudes twd losses	0.25***	1		
Gender	0.07***	0.05***	0.16***	
Age	0.16***	0.06***	0.27***	
Native	0.06***	0.03***	0.06***	
Paris	0.02***	0.06***	0.08***	
Matrimonial	0.10***	0.06***	0.22***	
Income	0.24***	0.14***	1	
Credit	0.10***	0.08***	0.41***	1
Stocks	0.30***	0.13***	0.18***	-0.03***
UL life insurance	0.32***	0.12***	0.12***	-0.02***
Retirement	0.09***	0.05***	0.07***	0.04***

Table 9 displays Spearman rank correlation coefficients corresponding to each binary/ordinal and ordinal/ordinal pairwise variables. T-test is used for assessing coefficients significance. Statistical significance levels are fixed at 1%, 5% and 10% and are denoted by ***, ** and *, respectively.

Table 9 shows that "Income" and "Credit" display a strong and positive association (r_{sp}=0.41) since banks only lend to clients whose income allows them to repay their debt. "Risk tolerance" is positively associated with "Stocks" (r_{sp}=0.30) and to "UL life insurance" (r_{sp}=0.32), both being two risky financial products. "Age" and "Income" are also positively associated reflecting wealth accumulation across time (r_{sp}=0.27). Furthermore, MiFID indicators are positively correlated with each other (r_{sp}=0.25). "Risk tolerance" and "Income" also display a positive and non-surprising association (r_{sp}=0.24). For "Attitudes twd losses", no high significant association is found with banking records variables.

B Variance Inflation Factor

Table 10 – Variance Inflation Factor (VIF)

	VIF		
	Model 1	**Model 2**	**Model 3**
	N=77,365	**N=71,461**	**N=71,745**
Panel A : MiFID indicators			
Risk tolerance			
0		(omitted)	
1		1.20	
2		1.03	
Attitudes twd losses			
1			1.05
2			1.02
3			(omitted)
4			1.03
Panel B : Socio-demographic indicators			
Gender	1.03	1.04	1.04
Age	2.94	2.88	2.88
Native	1.06	1.06	1.06
Paris	1.04	1.05	1.05
Matrimonial	1.07	1.07	1.07
Self-employed	1.15	1.14	1.14
Salaried	(omitted)	(omitted)	(omitted)
Retired	1.86	1.87	1.87
No occupation	2.30	2.22	2.23
Panel C : Wealth and patrimony indicators			
ln(Income)	2.60	2.54	2.53
ln(Credit)	1.37	1.36	1.36
UL life insurance	1.08	1.18	1.09
Retirement	1.03	1.03	1.03
Mean VIF	**1.54**	**1.48**	**1.43**

Table 10 reports VIF corresponding to independent variables. We do not face multicollinearity problem since all VIF are below the critical threshold of 10 (Chatterjee et al., 2000).

C *Goodness-of-fit measures*

We apply different measures for ensuring the quality of goodness-of-fit of our BLR and their degrees of prediction. We begin by interpreting statistical measures reported in Table 3 before conducting a thorough analysis.

In Table 3, we report the likelihood ratio chi-square (noted LR Chi2). This numerical measure of fit allows comparing the goodness-of-fit of each model to the intercept-only model, also called empty model. This first test shows that our models predict better than the empty model. We also report the p-value corresponding to each model. Whatever the model we focus on, p-value is always equal to 0.00 leading us to conclude that our models have a whole fit better than an empty model at all reasonable significance levels. Besides, pseudo-R^2 increases by adding MiFID variables into Model 1 (0.2028) leading thus to an increase of the quality of goodness-of-fit in models 2 (0.2446) and 3 (0.2102). Indeed, the closer the pseudo-R^2 is to 1, the stronger is the explanatory power of a model. Furthermore, by adding MiFID variables into Model 1, the log likelihood increases and then converges to 0 in Models 2 and 3 meaning that we converge to a good model (Cahuzac and Bontemps, 2008).

Table 11 reports four additional statistical measures.

Table 11 – Goodness-of-fit measures

Goodness-of-fit Measures	**Model 1 N=77,365**	**Model 2 N=71,461**	**Model 3 N=71,745**
AIC	42,906.74	39,182.34	41,055.16
BIC	43,027.08	39,319.99	41,202.06
Correct classification rate	89.23%	88.92%	88.66%
AUC	0.8186	0.8426	0.8215

Table 11 presents results obtained from goodness-of-fit measures corresponding to each model performed in Table 3. We report "Akaïke information criterion" (AIC), "Bayesian information criterion" (BIC), "Correct classification rate" and "Area Under the Curve" (AUC).

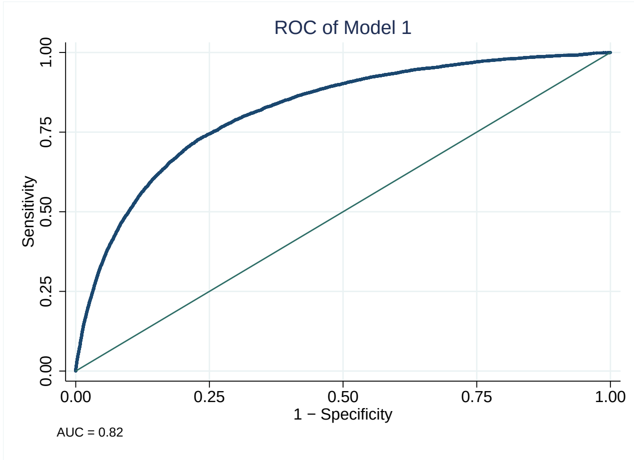
The selection criteria AIC and BIC of Models 2 and 3 are lower than those of Model 1. Therefore, the quality of Models 2 and 3 is higher than that of Model 1.

By comparing concordances and discrepancies between estimated values and observed values, we determine the explanatory power of all models, i.e. the correct classification rate. Specifically, we analyze our models' sensitivity and specificity. The sensitivity (specificity respectively) refers to the proportion of individuals who are positively (negatively respectively) and correctly classified. In our case, the sensitivity refers to the proportion of retail clients declared positive (i.e. stock holding) by the model and who are. The specificity refers to the probability of retail clients declared negative (non-stock holding) by the model and who are (i.e. they do not hold stock). Thereby, for detecting the correct classification rate, we must sum the number of retail clients (positively and negatively) correctly classified and divide it by the sample size. In our models, the correct classification rate is about 89% meaning that estimated values and observed values tie in 89%.

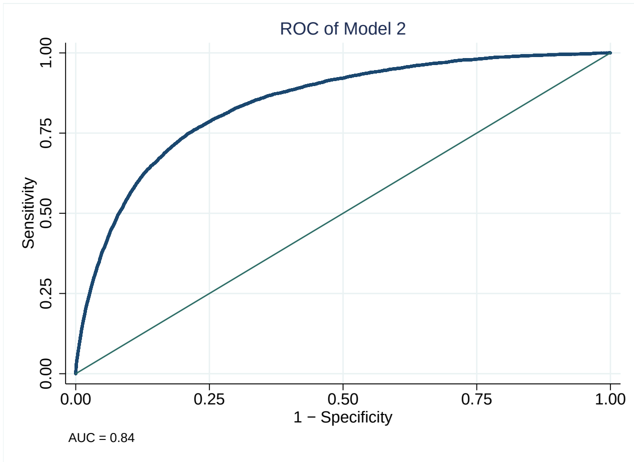
A graphical representation can also be realized for assessing the quality of a regression model. Indeed, we can use ROC (Receiver Operating Characteristics) curve for viewing the performance of a model. In other words, ROC represents graphically the discriminatory quality of a model. In our study, we aim at distinguishing retail clients holding stock from those who do not hold stock. For performing a ROC curve, we need to compute the sensitivity and the specificity. Graphically, “1 - specificity” is reported on the x-axis and “sensitivity” is reported on the y-axis. ROC curve provides a synthetic index, called Area Under the Curve (AUC). In our case, if AUC is equal to 1, then we can show that our model discriminates retail clients holding stock(s) from those who do not hold any stock in 100% of cases. Therefore, there is a strong discrimination. However, if AUC equals 0.5, this means that the probability of discriminating both retail clients is 50%. Therefore, the model is not informative since it is equivalent to a random selection. Graphically, an AUC equal to 0.50 represents a bisector. As its name suggests, we interest in the area comprised between ROC curve and the bisector. Then, the further away we locate from the bisector, the greater is the discriminatory quality of a regression. Therefore, we should obtain an AUC close to 1 for ensuring that the model discriminates stock holding from non-stock holding. In Table 11, we notice that the smallest AUC value is equal to 0.8186 (Model 1) and the highest one is equal to 0.8426 (Model 2). Graphical representation is reported in Figure 1. According to Long and Freese (2006), an AUC comprised between 0.80 and 0.90 means that there is a good discrimination. Therefore, we can conclude that all models reported in this study allow correctly discriminating retail clients holding stock(s) from those who do not hold any stock during the sample period. Specifically, Models 2 and 3 display higher AUC than that of Model 1 thus reinforcing the discrimination stockholding/non-stockholding.

Figure 1 – ROC and AUC

(a) Model 1



(b) Model 2



(c) Model 3

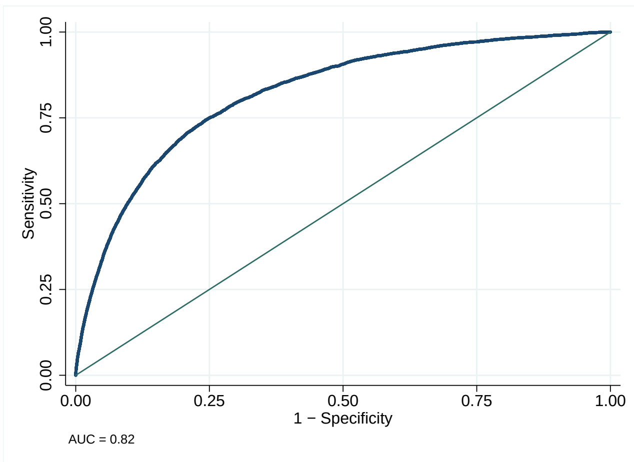


Figure 1 reports ROC and AUC corresponding to our models.

"Subjective Financial Literacy and Retail Investors' Behavior"

Bellofatto, Anthony ; D'Hondt, Catherine ; De Winne, Rudy

Abstract

This paper investigates the relationship between subjective financial literacy, i.e. self-reported by investors, and trading behavior. In particular, we use the level of financial knowledge and experience reported in the MiFID tests by retail investors. Such tests are implemented in the EU from the so-called MiFID directive since November 2007. We show that subjective financial literacy helps explain cross-sectional variations in retail investors' behavior. Investors who report higher levels of financial literacy seem to invest smarter, even after controlling for gender, age, portfolio value, trading experience and education. They trade more and are less prone to the disposition effect. They tend to concentrate their portfolios on a small set of stocks and achieve diversification through investment funds holding. Their trading behaviors allow them to display higher gross and net returns as well as higher excess Sharpe ratios. Our findings are relevant for both policy making and under...

Document type : *Document de travail (Working Paper)*

Référence bibliographique

Bellofatto, Anthony ; D'Hondt, Catherine ; De Winne, Rudy ; et. al. *Subjective Financial Literacy and Retail Investors' Behavior.* (2018)

Subjective Financial Literacy and Retail Investors' Behavior

Anthony Bellofatto, Rudy De Winne and Catherine D'Hondt*

Université catholique de Louvain - Louvain School of Management

This version: March 20, 2018

Abstract

This paper investigates the relationship between *subjective* financial literacy, i.e. self-reported by investors, and trading behavior. In particular, we use the level of financial knowledge and experience reported in the MiFID tests by retail investors. Such tests are implemented in the EU from the so-called MiFID directive since November 2007. We show that subjective financial literacy helps explain cross-sectional variations in retail investors' behavior. Investors who report higher levels of financial literacy seem to invest *smarter*, even after controlling for gender, age, portfolio value, trading experience and education. They trade more and are less prone to the disposition effect. They tend to concentrate their portfolios on a small set of stocks and achieve diversification through investment funds holding. Their trading behaviors allow them to display higher gross and net returns as well as higher excess Sharpe ratios. Our findings are relevant for both policy making and understanding retail investors' behavior.

JEL Classification: G02, G11, G28

Keywords : retail investors, financial literacy, MiFID

---

*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. They wish to thank C. Haddad, F. El Mouaaouy, M. Montone, V. Poti, N. Todorovic, A. Menkveld for helpful discussion and comments as well as participants at the 3L Finance Research Workshop, the 2015 Behavioural Finance Working Group Conference, the 2015 "EFM Merton H Miller" Doctoral Seminar of the European Financial Management Association, the 2015 Annual Meeting of the Academy of Behavioral Finance & Economics and the 2015 Southern Finance Association Annual Meeting. They also thank the participants at the LaRGE Research Center workshop in Strasbourg for their comments. Any errors are the full responsibility of the authors.
Contact address: Louvain Finance (IMMAQ), Louvain School of Management, Université catholique de Louvain - Chaussée de Binche 151, 7000 Mons, Belgium. Email adresses: anthony.bellofatto@uclouvain.be (corresponding author); rudy.dewinne@uclouvain.be; catherine.dhondt@uclouvain.be. Comments are welcome.

1 Introduction

Lusardi and Mitchell (2014) define *financial literacy* as the ability to process economic information and make informed decisions about financial planning, wealth accumulation, debt and pensions. In order to assess such an ability, these authors have designed a set of questions built on the three following basics: numeracy and capacity to do calculations related to interest rates, understanding of inflation, and understanding of risk diversification.[1] Their set of questions is now recognized as a standard in the literature. It has been administered to populations of different ages in the US (Lusardi and Mitchell (2011a)) but also in other countries such as the Netherlands (Van Rooij et al. (2011)) and Japan (Sekita (2011)). The main empirical findings all converge at a widespread low level of financial literacy. Beyond the level of individuals' knowledge of financial concepts, several authors show that financial literacy is effectively related to different aspects of financial behavior. For example, Hilgert et al. (2003) document a strong correlation between financial literacy and day-to-day financial management skills. In the same vein, Lusardi and Mitchell (2007) find that individuals with low financial literacy are less likely to plan for retirement and therefore accumulate less wealth during their lifetime. As for Guiso and Jappelli (2008), they provide evidence that measures of financial literacy are strongly correlated with the degree of portfolio diversification. Finally, a bunch of papers highlight the positive relationship between financial literacy and stock market participation (a.o. Kimball and Shumway (2006), Christelis et al. (2010), Van Rooij et al. (2011)).

Most of the above papers refer to *objective* measures of financial literacy, i.e. based on a set of questions designed to assess how people deal with fundamental concepts at the root of saving and investment decisions. Such objective measures reveal individuals' *actual* knowledge, where the latter is based on correct answers. By contrast, *subjective* measures of financial literacy rely on questions asking people to indicate their self-assessed financial knowledge and expertise. Such subjective data may best capture psychological drivers affecting the individual's decision-making process. Their use remains however quite infrequent in the financial literature, despite the growing amount of papers relying on surveys to elicit investors' attributes (a.o. Glaser and Weber (2007), Graham et al. (2009), Merkle and Weber (2014)). The reluctance towards such data is mainly an a priori skepticism: Can we trust what people state? Can we use this information to understand how they behave? And for financial literacy in particular,

[1] For more details, see Lusardi and Mitchell (2008), Lusardi and Mitchell (2011a) and Lusardi and Mitchell (2011b).

respondents are expected to be rather confident about their financial knowledge and, overall, overestimate how much they actually know. According to the literature, the relationship between *objective* and *subjective* literacy may not be taken for granted (Lusardi and Mitchell (2014)). While some authors document a strong positive relationship between both measures (Dorn and Huberman (2005), Van Rooij et al. (2011)), others report only a weak relationship (Guiso and Jappelli (2008), Lusardi (2011) and Bucher-Koenen et al. (2012)). Xia et al. (2014) even use the difference between both measures as a proxy of overconfidence.

In this paper, we investigate the relationship between subjective financial literacy and actual trading behavior. For that purpose, we use subjective measures of financial literacy available in the so-called MiFID tests. The latter are implemented in the EU since the MiFID Directive[2] came into force in November 2007. This piece of European regulation was wide and far reaching; it covered all forms of intermediation/services or dealing activities and impacted all financial intermediaries, their clients (either professional or retail) and the majority of financial instruments. In a nutshell, MiFID has made compulsory for investment firms to collect specific information about their retail clients' needs and preferences. Accordingly, investment firms operating in the EU are obliged to submit tests to their clients in order to determine their level of knowledge and experience, their investment objectives as well as their financial capacity. Such tests should help offer investors suitable services. Specifically, *suitability* assessment is required before providing investment advice or portfolio management services while *appropriateness* assessment is required before providing execution and transmission of orders (what is called "execution only" in the industry) on complex financial instruments. Basically, the *Suitability* test aims at understanding the types of investments that will be suitable for the investor while the *Appropriateness* test should assess the investor's knowledge and experience in complex financial instruments so as to protect those who would not understand or be

[2]MiFID stands for Markets in Financial Instruments Directive. Formerly known as Investment Services Directive II, this directive was the second step in the harmonization of the European capital markets industry. It essentially aimed at adapting the first Investment Services Directive (ISD 1, issued in 1993) to the realities of the current market structure. On October 2011, the European Commission adopted a legislative proposal for the revision of MiFID. This revision took the form of a revised Directive (MiFID II) and a new Regulation (MiFIR). In a nutshell, MiFID II came into force in January 2018 and confirms the role of the MiFID tests by strengthening conduct rules such as an extended scope for the Appropriateness test and reinforced information to clients. For more details, see the European Commission website (`http://ec.europa.eu/internal_market/securities/isd/mifid2/index_en.htm`.)

aware of the potential implications and level of risk involved in a "complex" transaction (i.e. involving "complex" instruments such as derivatives).

Although the MiFID tests have now been implemented for several years, they have raised little interest so far, whether in academia or in the financial industry. MiFID deserves though a particular attention since it requires investment firms to gather survey data about their clients but without defining standard questionnaires.[3] MiFID mainly requires that suitability assessment covers three sets of items: investment objectives, financial capacity, experience and knowledge. As for appropriateness assessment, it has to be based on experience and knowledge only. Furthermore, MiFID does not impose the use of objective measures of financial literacy and most of the time investors are rather asked to self-assess their level of financial knowledge. This wide latitude for interpretation has led to a large diversity of questionnaires since each investment firm has developed its own tests for profiling its clients.[4]

As for academic research addressing this topic, it is still in its infancy. Marinelli and Mazzoli (2011) document the differences characterizing the MiFID tests across 14 Italian investment firms. These authors show that the questionnaires largely diverge in their structure and content. According to them, this huge heterogeneity may have side effects leading to inconsistent profiling.[5] Linciano and Soccorso (2012) also analyze the questionnaires used by several Italian intermediaries and confirm that they depend on the firm's business model. These authors point out a lack of appropriate training courses for the advisors who have to administer these MiFID tests to clients.[6] Furthermore, they report that the tests under scrutiny mainly require self-assessment from clients and include several ambiguous questions that are easy to misunderstand. More recently, Mazzoli and Marinelli (2014) have focused on risk profiling for

[3]The European regulator only provided guidelines and general rules for implementing the MiFID tests.

[4]Supervisory authorities have taken initiatives to both evaluate how well the questionnaires used in practice comply with MiFID requirements and improve the implementing guidelines (a.o. AMF (2010), FSA (2011), ESMA (2012), FSMA (2014)). The resulting evidence tends to reveal the poor quality of suitability tests, the poor quality of client profiling, and poor advisory services as a consequence.

[5]They show that the same investor could be characterized as 'cautious' or 'dynamic', depending on the test that is used.

[6] While in a few cases the staff has been specifically trained, the training was limited to refresher courses on the legal aspects, or generic training courses for advisers. Workshops on the design of questionnaires were rarely included in the training sessions, nor explicit references to the potential issues of cognitive and behavioural biases affecting the administration of questionnaires. According to Linciano and Soccorso (2012), this represents a major issue since building valid and reliable questionnaires requires specific multidisciplinary skills.

a sample of 1,149 suitable portfolios and conclude that information gathered in the tests are not sufficient to determine an investor's risk profile.[7]

In contrast with the aforementioned papers, we aim at finding whether the answers given by retail investors in the MiFID tests are informative and consistent with their trading behavior. Specifically, we focus on financial literacy since it is included in both tests and should help investment firms elicit the degree of their clients' knowledge and experience. As such, this paper is, to the best of our knowledge, the first paper investigating the informativeness of financial literacy in the MiFID tests. Guiso and Jappelli (2008) document that "eliciting financial literacy by simply asking people if they know finance is bound to lead to serious mistakes [...] To put it simply, using self-assessment to rank investors on the basis of their financial knowledge for regulatory purposes is confounded by investors' over- or underconfidence." Our aim is therefore to determine whether the investors' self-assessment of their financial literacy may be useful for characterizing investors' trading behavior and may be reliable for both regulators and investments firms.

Our research question is relevant because the extant literature is still scarce and the results are often mixed. Dorn and Huberman (2005) are some of the first authors to confront investors' portfolios and trading activity with their own statements. They highlight that the inclusion of subjective investor attributes offers several insights into investor behavior. Regarding the relationship between self-assessment of financial literacy and trading behavior, they find ambiguous evidence. On the one hand, they report that investors who perceive themselves as more knowledgeable about financial securities display a better diversified portfolio. On the other hand, those who perceive themselves as better informed about financial securities than the average investor churn over their portfolios more, which may be evidence of overconfidence. Graham et al. (2009) focus on the "competence effect" and its impact on financial behavior.[8] They find that investors who feel competent trade more often and have more internationally diversified portfolios. Finally, Allgood and Walstad (2013) investigate the relationship between credit card behavior and financial literacy by using objective and subjective measures of literacy. Their results suggest that financial literacy is significantly related to less costly practices

[7] In particular, they put forward variables that are directly related to both the risk-holding and risk-allocation decisions of the Italian households in their sample.

[8] The competence effect could be related to the self-perceived financial literacy we analyze since it is defined as the fact of feeling skillful or knowledgeable in an area. The authors suggest that the competence effect is particularly relevant to investors' behavior as investors are constantly required to make decisions based on subjective probabilities.

in credit card use. However, they find perceived financial literacy to be a stronger predictor than actual financial literacy. This study also shows that the combination of subjective and objective assessments of financial literacy provides a more comprehensive analysis of how financial literacy affects credit card behavior. We should stress that none of the above papers dealt with the relationship between subjective literacy and performance.

In this paper, we use a unique database from an important online Belgian brokerage house to investigate the behavior of 20,285 retail investors during the 2003-2012 period. This database includes usual information relative to investors' orders and trades, but also their answers to both MiFID tests. Those tests are conducted online and answers are self-reported decisions the investors make on their own. The advantage is that the answers are not affected by any conversation with a broker or a financial advisor. However, online tests - like online trading activities in general - have also a drawback: they make investors "do-it-yourselfers" in an information-rich environment, thereby bolstering their overconfidence due to an illusion of both knowledge and control (Barber and Odean (2001), Volpe et al. (2002)).

In a first step, we check the consistency between subjective financial literacy reported in the Suitability test (hereafter S-test) and the one reported in the Appropriateness test (hereafter A-test). Like for any survey, the fact that such tests force investors to self-assess their financial literacy and report a lot of individual perceptions may raise skepticism about the meaningfulness of answers. In a second step, we check the consistency between subjective financial literacy and trading behavior characterized along three different aspects: experience and familiarity with financial markets, diversification and performance.

Our main findings may be summarized as follows. Regarding the self-assessment of financial literacy, our results show an overall consistency across investors' answers: investors who report a high literacy in the A-test are much more likely to also report a high literacy in the S-test. As for the consistency between subjective financial literacy and actual behavior, we provide empirical evidence that subjective literacy helps explain cross-sectional variations in retail investors' behavior. Investors who report higher levels of financial literacy tend to invest *smarter*. Specifically, they trade more on both stocks and complex instruments and they are less exposed to the disposition effect, which is consistent with higher experience. Although investors with higher subjective literacy trade in a larger stock universe, they hold less diversified stock portfolios (but not riskier). In fact, they tend to concentrate their stock portfolios on a small set of securities and achieve global diversification through investment funds holding. Finally, investors with higher subjective financial literacy display higher both gross and net returns as well as higher excess Sharpe ratios. All our results hold even when we control for gender, age, portfolio value, trading experience and education.

All in all, our findings support consistency between subjective literacy and actual trading behavior. Retail investors are overall consistent when reporting their financial literacy online. More importantly, this piece of information provided by the investors themselves could help better understand and characterize their actual trading behavior. Such results are relevant for both policy making and understanding retail investors' behavior. Subjective literacy reported in the MiFID tests is informative to characterize retail investors and hence deserve more attention in that perspective. This empirical evidence is meaningful for investment firms that are forced to administer the MiFID tests in the EU. Using subjective literacy could help those investment firms deliver the most suitable services to their retail clients. This paper could also provide insights for regulatory purposes, since we show that subjective financial literacy reported online does correlate with actual trading behavior. Generally speaking, our contribution to the literature appears even more relevant because the role of the MiFID tests has recently been confirmed in the European regulation.[9] It also opens new areas of research such as the role of opinions, perceptions and beliefs in the individuals' financial decision-making process.

The remainder of this paper is structured as follows. Section 2 describes our data and sample as well as the MiFID tests. We report our empirical work and the results in Section 3. Section 4 concludes.

2 Data and Sample

The data are provided by an online Belgian brokerage house and cover the period from January 2003 to March 2012. They refer to 20,285 retail investors and are made of two datasets. The first one contains information about the investors, that we classify into three categories. The first category includes socio-demographic data: year of birth and gender. The second category encompasses the answers to the A-test while the third category contains the answers to the S-test. The second dataset is made of detailed information about the investors' trading activity.

[9] With MiFID II that came into force in January 2018, the EU's regulation confirms the role of such tests by strengthening conduct rules such as an extended scope for the Appropriateness test and reinforced information to clients. For more details, see the European Commission website (`http://ec.europa.eu/internal_market/securities/isd/mifid2/index_en.htm`).

The online brokerage house provides their clients with an access to a large panel of financial instruments. The main traded securities are stocks, funds, options, warrants, and bonds. Only futures contracts cannot be directly traded on the common trading web platform of the broker.[10] The data include an ISIN code for each instrument, order size, price, type, executed quantity, trade price, time-stamps, explicit transaction costs as well as a code for the market where the trade was completed. Both datasets include an anonymized code identifying each investor, which allows us to select all information relative to a specific investor. For the purpose of our study, we use information about trading activity on stocks to build end-of-month portfolios for each investor in the sample. A third dataset including historical market data coming from Eurofidai and Bloomberg is then used to compute the market value of end-of-month stock portfolios.

2.1 Trading activity

Our sample contains 2,107,382 trades on stocks[11] executed by 20,285 investors across about 13,000 different stocks. 57% of the trades are purchases and 43% are sales on an aggregate basis. Individual investors execute about 235,000 trades in a typical year and about 20,000 trades in a typical month. Regarding socio-demographic data, the average investor is 48 years old[12] and we count only 10% of women in the sample. As for the level of education, 73% have a university degree, 20% hold a secondary/high school degree and 7% have no degree.

Tables 1 and 2 present descriptive statistics regarding trading activity. The average investor executes a total of 144 trades across all instruments. When focusing on stocks only, the average is 103 trades on 27 different stocks. With a sample period of 111 months, the average investor executes about 11 trades on stocks per year, which is similar to the trading activity reported in Kumar and Lee (2006).[13]

[10] As a result, we do not have data about the trading activity on futures contracts.

[11] We focus on stocks for which a valid ISIN code is available. For stocks traded in foreign currencies, we use exchange rates to convert monetary volumes into euros.

[12] Age is calculated in 2012 using the year of birth.

[13] Those figures are in line with other samples used in the literature on retail investors' trading activity. For example, Dhar and Zhu (2006) report an average of 60 trades on stocks per investor on a period from 1991 to 1995 (i.e. 12 trades per year per investor) and Barber and Odean (2000) find an average of 30 trades on stocks per investor over a 5-year period.

As for the other instruments, the average investor completes 26 trades on options or warrants, 12 trades on investment fund shares and less than 1 trade on bonds. The average trading experience is about 54 months (that is 4.5 years).[14] We observe a large dispersion in our sample. The number of total trades, the number of stock trades, the number of stocks traded and the trading experience are all positively skewed since the medians are substantially lower than the mean values. This positive skewness is even more striking for the number of trades on non-stock instruments. For example, only 25% of investors trade at least 2 options or 2 investment fund shares.

Table 1: Descriptive statistics for trading activity (1)

	Mean	Median	Q1	Q3
Number of total trades	144	54	19	144
Number of stock trades	103	40	15	102
Number of different stocks traded	27	16	7	34
Number of option trades	26.41	0	0	2
Number of fund trades	12.32	0	0	2
Number of bond trades	0.28	0	0	0
Trading experience (in months)	54	51	26	81

The table reports the cross-sectional mean, median, lower and upper quartiles for trade-based measures on a per investor basis over the sample period. 'Number of total trades' is the number of trades executed across all instruments. 'Number of stock trades' is the number of trades executed on stocks. 'Number of different stocks traded' is the number of different stocks traded during the whole trading period. 'Number of option trades' is the number of trades executed on both options and warrants. 'Number of fund trades' is the number of trades executed on investment fund shares. 'Number of bond trades' is the number of trades executed on bonds. 'Trading experience' is computed as the difference between the last trade date and the first trade date available in the sample. It is expressed in number of months.

Table 2 shows complementary statistics computed on binary variables. Regarding asset allocation, 34% of investors trade investment fund shares, 31% trade options or warrants, but only 8% trade bonds. These figures appear consistent both with the statistics reported in Table 1 and with papers dealing with other samples.[15]

[14]We compute the trading experience as the difference between the date of the last trade on stocks and the date of the first trade on stocks available in the sample. As in Glaser and Weber (2009), we exclude from our sample investors with less than 5 months of trading activity. This filter allows us to drop very short-lived investors.

[15]Boolell-Gunesh et al. (2012) report that about 12% of their French retail investors trade bonds and 25 % warrants. Korniotis and Kumar (2013) document that 22% of their sample investors hold mutual funds and 9%

Table 2: Descriptive statistics for trading activity (2)

	0	1
Bonds_trader	92%	8%
Funds_trader	66%	34%
Options_trader	69%	31%

The table reports statistics for trade-based measures built on binary variables. 'Bonds_trader', 'Funds_trader', 'Options_trader' are set to 1 when the investor made at least one trade on respectively bonds, investment fund shares and either options or warrants.

2.2 MiFID tests

MiFID came into force in 2007 across the EU member states. One of its objectives was to increase the level of protection of investment firms' clients. In addition to client categorization aiming at segregating retail investors from professional investors and eligible counterparts, MiFID requires investment firms to qualify their clients and the services requested through Suitability and Appropriateness tests. These two levels of qualification depend on the type of services provided to the investor.

The Suitability test (S-test) has to be submitted to investors before providing investment advice or portfolio management services. Assessment of suitability involves ensuring that the instruments and services offered meet the investor's objectives, financial capacity as well as his knowledge and experience in financial instruments. As mentioned earlier, the room for interpretation left by the regulator has generated a huge diversity of questionnaires used in the industry. In our case, the S-test under scrutiny is made of 11 questions. Among them, two questions directly deal with the level of knowledge of financial markets. For the purpose of our study, we only consider the answers to these two questions in the empirical part. Those are reported in Table 3, Panel B. Our sample is made of investors who asked for an access to an advice tool on stocks[16] and we have the S-test data for each of them.

trade at least once options. Koestner et al. (2017) find however a higher proportion of retail investors trading mutual funds in their sample (49%).

[16]During the sample period, the online brokerage house doesn't offer portfolio management services to its clients. It only provides a free access to an investment advice tool on stocks.

The Appropriateness test (A-test) has to be submitted to investors before providing execution and transmission of orders in complex financial instruments. Assessment of appropriateness mainly requires ensuring that the investor has the necessary experience and knowledge to understand the risks involved in complex financial instruments.[17] In practice, the brokerage house that provides us with data has implemented this test for an exhaustive list of instruments, including shares traded on a non-European market or on a European non-regulated market (such as Multilateral Trading Facilities under the MiFID typology). As a result, we have the answers to the A-test for all the retail investors of our sample so that it does not suffer from any selection bias. The A-test under scrutiny is made of 9 questions, among which one is about the general knowledge of financial markets. The answer to this specific question, which is provided in Panel A of Table 3, is considered in the empirical part.

[17] Unlike the S-test, the result of the A-test is not restrictive for investors. Based on the collected information, investment firms have only to provide recommendations about the appropriateness of financial instruments. Therefore an investor may choose, at his own risk, to go ahead with a transaction even if the involved instrument is flagged as inappropriate.

Table 3: Subjective financial literacy questions

Panel A: A-test question	Answers
What is your knowledge of financial markets?	
Level 0	I know a few things but I am interested by the financial markets
Level 1	I have sufficient experience to understand well the importance of a good diversification of risks
Level 2	I understand the functioning of the financial markets. I know 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
Level 3	I consider myself as an experienced investor who manages any aspect of the financial markets
Panel B: S-test questions	**Answers**
What is your knowledge of the financial markets?	
Level 1	I know very little about it and I am not really interested in it
Level 2	I am not familiar with investments, but I am interested in it
Level 3	I have sufficient experience to acknowledge the importance of risk diversification
Level 4	I have a good knowledge of the financial markets. I am aware that the financial markets can strongly fluctuate, that sector and asset categories have different characteristics regarding revenue, growth and risk profile
Level 5	I consider myself as an experienced investor who thoroughly masters all the aspects of the financial markets
How do you estimate your level of knowledge and experience about risks and potential obligations inherent to shares, bonds, funds and structured products?	(*based upon the type of product in which you have the lowest experience*)
Level 0	No knowledge
Level 1	Average knowledge
Level 2	Good knowledge

The table reports the questions dealing with subjective financial literacy in the MiFID tests under scrutiny. Panel A refers to the single question about financial knowledge in the A-test. Panel B presents the two questions about financial knowledge and experience in the S-test.

We should stress that the answers to both MiFID tests are online decisions made by investors themselves, without intermediaries. They are therefore not affected by conversations with a broker or a financial advisor. In addition, both tests include subjective literacy assessment but, as shown in Table 3, they do not ask exactly the same questions and available answers differ. Investors are not necessarily forced to fill in both tests at the same time. One shortcoming of our data is that they report neither the date at which the investor filled in the tests nor their potential updates.

Statistics for subjective literacy are provided in Table 4. In Panel A, we observe a large heterogeneity among investors. Only 11% of investors consider themselves as experienced investors while 20% of them report a basic knowledge. The most frequently chosen level is the third on the scale, 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'.

In Panel B, the empirical frequencies regarding the 'knowledge of financial markets' seem to be somewhat consistent with those observed in Panel A despite the use of a different scale. 9% of investors view themselves as very experienced investors in the S-test but only 3% report that they know very little about financial markets. The first two levels represent about 17% of investors. Again, the most frequently selected level is the second last on the scale, which states that the investor 'has a good knowledge of the financial markets and is aware that the financial markets can strongly fluctuate, that sector and asset categories have different characteristics regarding revenue, growth and risk profile'. The second question in Panel B covers both the knowledge and experience about "complex" instruments.[18] 56% (30%) of investors consider they have an average (a good) knowledge and experience. Only a minority (14%) of investors report they have no knowledge and experience.

[18]As described in Table 3, investors are asked to answer this specific question based upon the type of product in which they have the lowest experience. The listed instruments are shares, bonds, funds and structured products.

Table 4: Statistics for questions about subjective financial literacy

Panel A: A-test question					
	0	1	2	3	
Knowledge of financial markets	20%	28%	41%	11%	
Panel B: S-test questions					
	1	2	3	4	5
Knowledge of financial markets	3%	14%	31%	43%	9%
	0	1	2		
Knowledge and experience about "complex" instruments	14%	56%	30%		

The table reports the empirical frequencies for subjective financial literacy in the MiFID tests. In Panel A, levels 0 to 3 refer to the answers to the question about financial knowledge in the A-test as reported in Table 3. In Panel B, levels 1 to 5 refer to the answers to the question about financial knowledge in the S-test as reported in Table 3. Levels 0 to 2 refer to the answers to the question about knowledge and experience about "complex" instruments in the S-test as reported in Table 3.

2.3 Stock portfolio data

As mentioned earlier, we use data on stock-related 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. Combining our data with historical market data, we also compute the monthly average portfolio value as well as the monthly average turnover as in Hoffmann et al. (2013).[19] Building on the literature, we assume that all trades executed in a given month take place on the last day of this month to finally compute both gross and net monthly returns.[20]

Table 5 reports descriptive statistics for the above-mentioned measures. From Panel A, we know that the average investor holds a six-stock portfolio while the median investor holds a four-stock portfolio. The average end-of-month portfolio value is about € 44,000 with a median value of about € 11,000. As for the turnover, the average investor churns 0.285 times

[19] Average of the absolute values of all purchases and sales in a particular month divided by the average of the portfolio values at the beginning and the end of this particular month.

[20] This assumption is used in Barber and Odean (2000), Barber and Odean (2001), Shu et al. (2004) and Glaser and Weber (2007). Barber and Odean (2000) show that this simplification does not bias the measurement of portfolio performance.

his portfolio each month (with a median of 0.1053). These figures are overall in line with other papers dealing with retail investors' portfolios (a.o. Barber and Odean (2001), Shu et al. (2004), Dorn and Huberman (2005), Kumar and Lee (2006), Glaser and Weber (2007), Goetzmann and Kumar (2008), Hoffmann et al. (2013)). Like the trade-based variables, all these portfolio-based variables are positively skewed since the means are substantially larger than the medians.

In Panel B, we report descriptive statistics for the average monthly gross and net return per investor, using both an arithmetic and a geometric average. Although the means of the arithmetic (gross and net) returns are quite large,[21] the median value is consistent with other empirical evidence (a.o. Glaser and Weber (2007)). The mean of the geometric gross returns is equal to 0.6% while the mean of the geometric net returns is not statistically different from zero. The average volatility of monthly gross returns is 28% while the median is 12%. These performance measures also display a positive skewness and reveal a large heterogeneity in our sample. This is consistent with Barber and Odean (2013) who point out that the aggregate performance of retail investors masks tremendous variations in behavior and in outcomes across individuals.

[21] Dorn and Huberman (2005) report quite similar figures, with a mean value of 2%.

Table 5: Descriptive statistics for end-of-month portfolio data

	Mean	Median	Q1	Q3
Panel A: Monthly stock portfolio				
Number of stocks	6	4	2	8
Portfolio value (€)	43,844	11,136	3,513	32,755
Turnover (%)	28.5	10.53	5.64	22
Panel B: Monthly trading performance				
Arithmetic gross return (%)	2.62	0.91	-0.32	3.13
Arithmetic net return (%)	1.64	0.48	-0.84	2.58
Geometric gross return (%)	0.6	0.38	-0.8	1.85
Geometric net return (%)	0	0	-1.3	1.44
Volatility (%)	27.87	12	8	22

The table reports the cross-sectional mean, median, lower and upper quartiles for portfolio-based measures on a per investor basis over the sample period. 'Number of stocks' is the monthly average number of stocks held in portfolio. 'Portfolio value' is the monthly average portfolio market value. 'Turnover' is the monthly average turnover. It is calculated as in Hoffmann et al. (2013), i.e. average of the absolute values of all purchases and sales in a particular month divided by the average of the portfolio values at the beginning and the end of this particular month. 'Arithmetic gross return' is the arithmetic average of monthly gross returns. 'Arithmetic net return' is the arithmetic average of monthly net returns. 'Geometric gross return' is the geometric average of monthly gross returns. 'Geometric net return' is the geometric average of monthly net returns. 'Volatility' is the standard deviation of monthly gross returns.

2.4 Measures of trading behavior and financial literacy

Some of the above variables will be used in the empirical part to characterize investors' trading behavior along three different aspects: experience and familiarity with financial markets, diversification and performance.

As measures of experience and familiarity with financial markets, we use the number of total trades across instruments, the number of stock trades and the monthly average turnover. Building on Goetzmann and Kumar (2008), we also consider whether investors trade options or warrants since these derivative securities create high entry barriers that individuals with low financial literacy may find difficult to overcome.[22] Moreover, we consider the retail investors' exposure to the Disposition Effect (DE hereafter), which refers to investors' reluctance to

[22] For example, Hsiao and Tsai (2018) provide evidence that individual investors with higher levels of knowledge are more likely to trade derivatives.

realize losses (i.e. they keep “losers”) as well as their propensity to realize gains (i.e. they sell “winners”). This behavioral bias, which was first labelled by Shefrin and Statman (1984), is today well-documented and several papers show that investors' experience helps dampen it.[23] We apply the methodology of Odean (1998) to assess the DE at the individual level, i.e we measure this bias as the difference between the proportion of gains realized and the proportion of losses realized.

To assess diversification we use the number of different stocks traded during the whole period since it reveals how large an investor' stock investment universe is. In addition, we use the monthly average number of stocks held in portfolio. While Goetzmann and Kumar (2008) state that the number of stocks in a portfolio is a useful heuristic for identifying the degree of diversification, other authors report that this “crude” measure of diversification often overstates the actual level of diversification (a.o. Blume and Friend (1975)). Therefore, building on Dorn and Huberman (2005), we use the volatility of monthly returns as a complementary measure of risk diversification and we also consider whether investors trade investment funds. According to Guiso and Jappelli (2008), diversifying wealth through funds requires a good understanding of the diversification benefits as well as the risk properties of the assets pooled within the fund. As in Goetzmann and Kumar (2008) and Koestner et al. (2017), we finally add the monthly average Herfindahl-Hirschman Index (HHI hereafter). This index of diversification can be approximated by the sum of squared stock portfolio weights. HHI ranges then from 0 (for well-diversified portfolios) to 1 (for underdiversified portfolios including only one stock). However, for investors who hold monthly positions in investment funds, we adjust the index by replacing any position in funds by a portfolio of 50 equally-weighted securities as in Dorn and Huberman (2005) and Koestner et al. (2017). In that case, we refer to it as the modified HHI (M_HHI hereafter).

[23]Feng and Seasholes (2005) show that retail investors with higher trading experience display a lower DE. Dhar and Zhu (2006) provide evidence that the exposure to the DE is lower among investors who are older, have a higher professional status and larger income. Boolell-Gunesh et al. (2012) find that investors who trade derivatives, bonds, and hold multiple accounts to place orders are less exposed to the DE.

As performance measures, we first use the geometric average of both gross and net returns.[24] We also consider Sharpe ratios that are risk-adjusted measures.[25] In addition, to take into account the market performance, we compute excess Sharpe ratios, i.e. Sharpe ratios in excess of market Sharpe ratio. The latter is then measured as the monthly market portfolio return in excess of the risk-free rate compared to the volatility of monthly market portfolio returns.[26] For each investor, we end up then with six measures of performance that are his monthly average gross and net returns, Sharpe ratio and excess Sharpe ratio.

In the empirical part, we will relate the above measures that characterize investors' actual behavior with measures of subjective literacy. For the latter, we will directly use the questions from the MiFID tests presented in Table 3.

3 Empirical work

3.1 Consistency across investors' answers in both MiFID tests

Our motivation to assess the consistency across investors' answers for similar questions in both MiFID tests is threefold. First, both the S-test and the A-test force investors to self-assess their financial literacy, which may cast doubt about the meaningfulness of answers. Second, investors are not forced to fill in both tests at the same time because they depend on different services. Third, if both tests include literacy assessment, they do not ask exactly the same questions. The ordering of questions, the wording of questions or even the scales presented to investors could activate cognitive factors that affect the way they assess their knowledge (Bertrand and Mullainathan (2001)). These three phenomena could lead to some inconsistency across the answers provided by the same investor, an effect that Tversky and Kahneman (1981) name the "framing effect".[27]

[24] Given the high volatility of monthly returns over the sample period, the arithmetic average of monthly returns is not a representative measure of performance.

[25] Sharpe ratio measures the portfolio return in excess of the risk-free rate compared to the portfolio's riskiness as measured by the volatility of portfolio returns. In this paper, we use the monthly-equivalent 12-month Belgian T-bill rate as a proxy for the risk-free rate.

[26] As a benchmark for the market portfolio, we use the Eurostoxx 600 Index.

[27] Bruine de Bruin (2011) states that the framing effect may be due to variations in question wording, choice set, and presentation order.

For our purpose, we use contingency tables wherein unconditional and conditional empirical frequencies are reported. We focus on self-reported financial knowledge in both MiFID tests and provide the results in Table 6. Based on the χ^2 statistic,[28] we first strongly reject the null hypothesis of independence between the two measures of subjective financial literacy. Comparing unconditional to conditional frequencies, an investor who reports a high level of literacy in the A-test is much more likely to mention a high level of financial knowledge in the S-test. For example, while the unconditional empirical frequency for the investors who choose the highest level of literacy in the S-test is about 9%, the corresponding frequency increases to about 48% for the investors who also select the highest level of knowledge in the A-test.

Then, in order to assess the level of consistency between the two measures of subjective financial literacy, we use the Spearman's rank correlation. It exhibits a value of 54%, thereby confirming a high but not perfect consistency across answers. Such a perfect correlation should not be realistic because of the difference of scale, the difference of wording as well as the potential difference of timing between both tests.

Finally, the number of investors who provide totally inconsistent answers is substantially low. Only 66 (80) investors have selected the highest (lowest) level of knowledge in the A-test while they have chosen the lowest (highest) level of financial knowledge in the S-test, accounting for only 0.72% of investors in our sample.

---

[28] $\chi^2 = \sum_i \sum_j \frac{(n_{ij} - \frac{n_i n_j}{n})^2}{\frac{n_i n_j}{n}}$ where the degree of freedom is $(r-1)(c-1)$ with r the number of rows and c the number of columns.

Table 6: Subjective financial literacy in the A-test vs. in the S-test (1)

A-test		S-test					
		1	2	3	4	5	Total
0	(#)	166	1,847	1,172	797	80	4,062
	(%)	0.82	9.11	5.78	3.93	0.39	20.02
	(r%)	4.09	45.47	28.85	19.62	1.97	
	(c%)	28.92	64.81	18.85	9.04	4.39	
1	(#)	188	745	2,898	1,765	106	5,102
	(%)	0.93	3.67	14.29	8.70	0.52	28.11
	(r%)	3.30	13.07	50.82	30.95	1.86	
	(c%)	32.75	26.14	46.61	20.01	5.81	
2	(#)	154	234	1,952	5,379	547	8,266
	(%)	0.76	1.15	9.62	26.52	2.70	40.75
	(r%)	1.86	2.83	23.61	65.07	6.62	
	(c%)	26.83	8.21	31.40	60.98	30.01	
3	(#)	66	24	195	880	1,090	2,255
	(%)	0.33	0.12	0.96	4.34	5.37	11.12
	(r%)	2.93	1.06	8.65	39.02	48.34	
	(c%)	11.50	0.84	3.14	9.98	59.79	
Total	(#)	574	2,850	6,217	8,821	1,823	20,285
	(%)	2.83	14.05	30.65	43.49	8.99	100.00

Statistics	Value	P-value
χ^2	11,291	<.0001
Spearman's rank correlation	0.54	<.0001

This contingency table reports respectively, for each pair of answers, the empirical frequencies (#), the total percentages (%), the row percentages (r%) and the column percentages (c%). Answers for the A-test are positioned in rows while those for the S-test are in columns. For the question in the A-test, levels 0 to 3 refer to the answers to the question about financial knowledge in the A-test as reported in Table 3. For the question in the S-test, levels 1 to 5 refer to the answers to the question about financial knowledge in the S-test as reported in Table 3. The results for the Chi-Square test for the null hypothesis of independence are also provided as well as the Spearman's rank correlation.

As a robustness check, we replicate the same analysis with another combination of similar questions: the self-reported financial knowledge in the A-test and the self-reported knowledge and experience about "complex" instruments in the S-test. Our findings are still consistent and support overall consistency across investors' answers in both tests. Table 12 in appendix exhibits the results.

3.2 Consistency between subjective financial literacy and trading behavior

We now investigate the relationship between subjective financial literacy and trading behavior. For that purpose, we focus on two measures of subjective financial literacy that are the self-reported level of financial knowledge in the A-test and the self-reported level of knowledge and experience about "complex" instruments in the S-test.[29] The variables presented in Subsection 2.4 are used to characterize investors' trading behavior. We still distinguish measures of experience and familiarity with financial markets, diversification and performance.

3.2.1 Univariate analysis

In a preliminary step, we perform an analysis of variance (ANOVA hereafter) to investigate whether our measures characterizing trading behavior significantly vary across the different levels of subjective literacy. Table 7 reports the results for the levels of financial knowledge in the A-test while Table 8 provides the results for the levels of knowledge and experience about "complex" instruments in the S-test. In line with the consistency highlighted in Subsection 3.1, Tables 7 and 8 display similar findings.

All ANOVA results exhibit highly significant F-stat values, except for the monthly turnover. Our findings suggest a positive relationship between investors' experience and familiarity with financial markets and their subjective financial literacy. Specifically, investors who assess themselves as highly literate tend to display a higher trading activity on stocks and other instruments. Similarly, the proportion of option traders is higher among investors who choose the highest levels of financial literacy. Moreover, all categories of investors display a positive DE but the latter is lower for investors who report higher levels of financial literacy.

When focusing on diversification measures, the results suggest that investors who report a higher level of financial literacy tend to trade in a larger universe of stocks, to hold a higher number of stocks in portfolio, and to be more active on investment funds, which can be associated with a higher awareness of the diversification concept. In addition, the HHI is overall lower for investors who report higher levels of financial literacy, thereby suggesting they hold better diversified stock portfolios. When taking into account the monthly holding in

[29]The results for the self-reported financial knowledge in the S-test are qualitatively similar and are available upon request.

funds in the modified HHI, this finding is even more striking. However, the results also show higher levels of stock portfolio volatility for investors with higher literacy.

As for performance, investors who report a higher level of financial literacy display significantly higher gross and net returns. It is worth to point out that only investors who choose the highest level of subjective financial literacy earn on average positive net returns. The relationship still holds on a risk-adjusted basis since both gross and net Sharpe ratios increase along with levels of literacy. Moreover, investors who perceived themselves as highly literate tend to display positive gross excess Sharpe ratios. Nevertheless, when taking into account explicit transaction costs, only those who select the highest level of financial literacy in the A-test are able to perform as well as the market portfolio on a risk-adjusted basis.

Table 7: ANOVA results for subjective financial literacy in the A-test and trading behavior

		Financial knowledge				
		0	1	2	3	F-stat
Experience and familiarity	Number of total trades	74.11	110.27	168.77	272.22	167.02***
	Number of stock trades	60.34	86.62	118.83	171.20	84.66***
	Turnover (%)	29.17	27.60	27.38	33.82	0.93
	Option_trader (%)	13.98	20.69	37.45	63.76	792.00***
	DE	13.94	12.85	10.94	9.90	22.91***
Diversification	Number of different stocks traded	16.63	23.91	31.71	39.79	270.22***
	Number of stocks	4.41	5.64	6.91	7.30	127.97***
	Volatility (%)	24.78	27.49	28.83	30.83	5.33***
	Fund_trader (%)	21.07	29.12	40.78	47.93	249.71***
	HHI	0.55	0.48	0.45	0.48	113.84***
	M_HHI	0.51	0.44	0.39	0.41	175.80***
Performance	Gross return (%)	0.28	0.53	0.80	0.92	23.43***
	Net return (%)	-0.43	-0.06	0.18	0.22	24.42***
	Gross Sharpe-ratio (%)	-0.88	-0.93	0.10 (NS)	0.56	3.07**
	Net Sharpe-ratio (%)	-5.15	-4.55	-3.72	-3.09	3.99***
	Gross excess Sharpe-ratio (%)	-1.31	0.42 (NS)	2.35	3.77	24.17***
	Net excess Sharpe-ratio (%)	-5.58	-3.20	-1.47	0.11 (NS)	26.48***

The table reports the results for the analysis of variance (ANOVA) on the relationship between several variables characterizing trading behavior and subjective financial literacy in the A-test. For each variable under scrutiny, we report its mean for each level of literacy. Levels 0 to 3 refer to the available answers to the specific question about financial knowledge in the A-test as described in Table 3. 'Number of total trades' is the number of trades executed across all instruments. 'Number of stock trades' is the number of trades executed on stocks. 'Turnover' is the monthly average turnover, expressed in %. It is calculated as in Hoffmann et al. (2013), i.e. average of the absolute values of all purchases and sales in a particular month divided by the average of the portfolio values at the beginning and the end of this particular month. 'Option_trader' is the proportion of investors who made at least one trade on either options or warrants. 'DE' refers to the disposition effect computed at the individual level as the difference between the proportion of gains realized and the proportion of losses realized. 'Number of different stocks traded' is the number of different stocks traded during the whole trading period. 'Number of stocks' is the monthly average number of stocks held in portfolio. 'Volatility' is the standard deviation of the stock portfolio monthly gross returns. 'Fund_trader' is the proportion of investors who made at least one trade on investment funds. 'HHI' is the monthly average Herfindahl-Hirschman Index, which is computed as the sum of squared stock portfolio weights. 'M_HHI' is a modified version of the HHI for which any position in funds is replaced by a portfolio of 50 equally-weighted securities. 'Gross return' is the geometric average of monthly gross returns. 'Net return' is the geometric average of monthly net returns. 'Gross Sharpe-ratio' is a risk-adjusted measure of gross return. 'Net Sharpe-ratio' is a risk-adjusted measure of net return. 'Gross excess Sharpe-ratio' is the gross Sharpe ratio in excess of market Sharpe ratio. 'Net excess Sharpe-ratio' is the net Sharpe ratio in excess of market Sharpe ratio. The last column exhibits the F-stat values as well as their significance. *** indicates significance at 1%; ** indicates significance at 5%; * indicates significance at 10%. 'NS' in brackets indicates that the mean value is not statistically different from zero.

Table 8: ANOVA results for subjective financial literacy in the S-test and trading behavior

		Knowledge and experience about “complex” instruments			
		0	1	2	F-stat
Experience and familiarity	Number of total trades	80.21	119.05	224.11	210.72***
	Number of stock trades	61.39	91.99	146.49	105.51***
	Turnover (%)	31.53	27.06	29.79	1.04
	Option_trader (%)	15.64	24.31	50.74	905.81***
	DE	13.88	12.45	10.15	29.64***
Diversification	Number of different stocks traded	17.78	25.03	36.43	319.41***
	Number of stocks	4.62	5.87	7.23	130.74***
	Volatility (%)	23.22	27.72	30.36	12.05***
	Fund_trader (%)	22.44	31.13	45.84	301.77***
	HHI	0.54	0.48	0.46	84.37***
	M_HHI	0.50	0.43	0.39	147.32***
Performance	Gross return (%)	0.34	0.56	0.91	28.74***
	Net return (%)	-0.33	-0.07	0.28	26.99***
	Gross Sharpe-ratio (%)	-0.78	-0.76	0.67	6.06***
	Net Sharpe-ratio (%)	-5.05	-4.47	-3.18	6.12***
	Gross excess Sharpe-ratio (%)	-0.23 (NS)	0.70	2.91	17.53***
	Net excess Sharpe-ratio (%)	-4.5	-3.00	-0.93	17.73***

The table reports the results for the analysis of variance (ANOVA) on the relationship between several variables characterizing trading behavior and subjective financial literacy in the S-test. For each variable under scrutiny, we report its mean for each level of literacy. Levels 0 to 2 refer to the available answers to the specific question about knowledge and experience about “complex” instruments in the S-test as described in Table 3. ‘Number of total trades’ is the number of trades executed across all instruments. ‘Number of stock trades’ is the number of trades executed on stocks. ‘Turnover’ is the monthly average turnover, expressed in %. It is calculated as in Hoffmann et al. (2013), i.e. average of the absolute values of all purchases and sales in a particular month divided by the average of the portfolio values at the beginning and the end of this particular month. ‘Option_trader’ is the proportion of investors who made at least one trade on either options or warrants. ‘DE’ refers to the DE computed at the individual level as the difference between the proportion of gains realized and the proportion of losses realized. ‘Number of different stocks traded’ is the number of different stocks traded during the whole trading period. ‘Number of stocks’ is the monthly average number of stocks held in portfolio. ‘Volatility’ is the standard deviation of the stock portfolio monthly gross returns. ‘Fund_trader’ is the proportion of investors who made at least one trade on investment funds. ‘HHI’ is the monthly average Herfindahl-Hirschman Index, which is computed as the sum of squared stock portfolio weights. ‘M_HHI’ is a modified version of the HHI for which any position in funds is replaced by a portfolio of 50 equally-weighted securities. ‘Gross return’ is the geometric average of monthly gross returns. ‘Net return’ is the geometric average of monthly net returns. ‘Gross Sharpe-ratio’ is a risk-adjusted measure of gross return. ‘Net Sharpe-ratio’ is a risk-adjusted measure of net return. ‘Gross excess Sharpe-ratio’ is the gross Sharpe ratio in excess of market Sharpe ratio. ‘Net excess Sharpe-ratio’ is the net Sharpe ratio in excess of market Sharpe ratio. The last column exhibits the F-stat values as well as their significance. *** indicates significance at 1%; ** indicates significance at 5%; * indicates significance at 10%. ‘NS’ in brackets indicates that the mean value is not statistically different from zero.

3.2.2 Multivariate analysis

Trading behavior may differ across investors who report high and low levels of financial knowledge because subjective literacy correlates with other attributes that predict trading behavior. In the literature, gender, age and income are recognized as some of the major drivers of trading behavior (e.g. Barber and Odean (2001), Goetzmann and Kumar (2008), Hackethal et al. (2012), Graham et al. (2009), Hoffmann et al. (2013)). Furthermore, Lusardi and Mitchell (2014) report that financial knowledge significantly depends on the level of education.[30] In the same vein, several papers document that education is significantly related to financial behavior (a.o. Haliassos and Bertaut (1995) and Campbell (2006)). However, controlling for the level of education does not decrease the impact of financial literacy but it can even enhance it as shown in Lusardi and Mitchell (2011a) and Van Rooij et al. (2011). Lusardi and Mitchell (2014) conclude therefore that general knowledge and financial knowledge both contribute to explain financial behavior.

In order to assess whether different trading behaviors can be related to differences in subjective financial literacy, we estimate cross-sectional regressions wherein the dependent variables are our measures characterizing trading behavior (see Subsection 2.4) and the set of explanatory variables includes several dummies based on subjective literacy[31] as well as control variables such as age, gender, income (that we proxy by the natural logarithm of the monthly average portfolio value) and the level of education. We also control for trading experience, i.e. the number of months during which an investor actively trade within the sample period.

Tables 9, 10 and 11 report the results for the regressions including three dummies for the three highest levels of subjective literacy in the A-test.[32] For continuous dependent variables, parameters are estimated thanks to OLS regressions while Logit models are used for binary dependent variables.[33]

[30]Lusardi and Mitchell (2007) and Lusardi and Mitchell (2011b) bring evidence that individuals without a college education are less likely to know basic financial literacy concepts.

[31]For each MiFID question, we include N-1 dummies and consider the lowest level as the category of reference.

[32]Tables 13, 14 and 15 in appendix report the results for the regressions including two dummies for the two highest levels of knowledge and experience in "complex" instruments in the S-test. The results are qualitatively similar.

[33]Building on Glaser and Weber (2007), we use the natural logarithm of the number of total trades across instruments, the number of stock trades, the turnover, the number of different stocks traded, the average number of stocks held in portfolio, the volatility, the HHI and Modified HHI since these variables are positively skewed. The authors state that this methodology allows to avoid problems of normality, nonlinearity and heteroscedas-

In Table 9, we focus on measures of experience and familiarity with financial markets. The results show that investors who perceive themselves as highly literate display a higher trading activity (whatever the instruments), churn over their portfolios more, and are also more likely to invest in options or warrants. This suggests an obvious higher level of experience and familiarity with financial markets. In addition, those investors tend to be less exposed to the DE, which is again consistent with a higher level of experience (Feng and Seasholes (2005), Dhar and Zhu (2006) and Boolell-Gunesh et al. (2012)).

Table 10 provides the results for our measures of diversification. Regression (1) brings evidence that investors who perceive themselves as highly literate trade in a larger stock universe while Regressions (2) and (3) suggest that those investors hold less diversified *stock* portfolios. However, Regression (4) shows that they do not hold riskier portfolios in terms of volatility. Furthermore, their higher tendency to invest in investment funds enables them to hold better diversified *global* portfolios. Hence, investors who report higher levels of financial literacy hold portfolios for which the modified HHI is lower. Taken all together, these findings suggest therefore that investors who report higher levels of financial literacy concentrate their stock portfolios on a small set of securities and achieve global diversification through investment funds holding.

The dependent variables are performance measures in Table 11. The results show that investors who report a higher level of financial literacy tend to display higher gross and net returns. However, this relationship does not hold anymore when focusing on Sharpe ratios. Regressions (5) and (6) suggest nevertheless that investors who perceive themselves as highly literate exhibit higher excess Sharpe ratios.

In Tables 9, 10 and 11, the results for our control variables are overall in line with the extant literature. They bring evidence that masculinity is positively related to trading activity while this attribute is negatively related to performance (Barber and Odean (2001)). As documented in Barber and Odean (2001) and Dorn and Huberman (2005), older investors tend to churn over their portfolios less and are less likely to invest in options and warrants. In addition, older investors are less exposed to the DE, are more prone to hold better diversified portfolios (as in Dorn and Huberman (2005) and Goetzmann and Kumar (2008)), and earn higher returns (as shown in Barber and Odean (2001)). As for the impact of income on trading behavior,

ticity in cross-sectional regressions. For the turnover, we compute the natural logarithm of (1+turnover) as in Glaser and Weber (2009) since a very low proportion of investors display an average turnover of 0. These investors typically build their portfolios during their first month of trading and do not change it afterwards.

our results show that investors who hold larger portfolios display a higher trading activity (as highlighted in Glaser (2003), Vissing-Jorgensen (2003) and Abreu and Mendes (2012)), are less exposed to the DE, hold better diversified portfolios (as shown in Dorn and Huberman (2005), Guiso and Jappelli (2008) and Goetzmann and Kumar (2008)), and also earn higher returns. Only our results about education are mixed. We find that investors with a university degree tend to churn over their portfolios less, exhibit a lower DE, hold more diversified portfolios but do not earn higher returns.

Table 9: Results for subjective financial literacy in the A-test & experience and familiarity with markets

	(1) Ln(total_trades)	(2) Ln(stock_trades)	(3) Ln(1+turnover)	(4) O_trader	(5) DE
Intercept	0.03	-0.60***	2.53***	-3.51***	30.76***
Gender	0.09***	0.06***	0.10***	0.03	-1.43**
Age	0.01	-0.01	-0.01***	-0.01***	-0.17***
Level of education 1	0.18***	0.17***	0.11***	0.08	-0.28
Level of education 2	-0.01	-0.04	-0.13***	0.08	-1.68**
Ln(PF value)	0.29***	0.36***	0.05***	0.16***	-0.81***
Trading experience	0.02***	0.02***	-0.01***	0.01***	0.01
Financial markets knowledge 1	0.06***	-0.01	-0.01	0.27***	0.21
Financial markets knowledge 2	0.23***	0.04**	0.04*	0.98***	-1.07**
Financial markets knowledge 3	0.55***	0.14***	0.21***	2.04***	-2.19***
Adjusted R^2	44.26%	54.21%	2.73%	-	1.81%
Pseudo R^2	-	-	-	12.76%	-
N	20,285	20,285	20,285	20,285	20,285

The table reports the regression results for the relationship between our measures of experience and familiarity with financial markets and subjective financial literacy in the A-test. The dependent variables of Regressions (1) to (5) are the natural logarithm of the total number of trades across instruments, the natural logarithm of the number of stock trades, the natural logarithm of (1+turnover), a binary variable set to 1 when the investor traded at least once options or warrants, and the investor's DE measured as the difference between the proportion of gains realized and the proportion of losses realized. In the set of explanatory variables, 'Gender' is a dummy set to 1 for males, 'Age' is the investor's age in 2012, 'Level of education 1' is a dummy set to 1 for investors with a secondary/high school degree, 'Level of education 2' is a dummy set to 1 for investors with a university degree, 'Ln(PF value)' is the natural logarithm of the average monthly portfolio value (as a proxy of wealth) and 'Trading experience' is defined as the difference between the date of the last trade and the date of the first trade on stocks. The last three variables 'Financial markets knowledge 1', 'Financial markets knowledge 2', 'Financial markets knowledge 3' are dummies used respectively for the levels 1, 2 and 3 in the question about financial knowledge in the A-test as reported in Table 3. *** indicates significance at 1%; ** indicates significance at 5%; * indicates significance at 10%. 'N' gives the number of observations used in each model.

Table 10: Results for subjective financial literacy in the A-test & Diversification

	(1) Ln(n_stocks)	(2) Ln(n_stocks_PF)	(3) Ln(HHI)	(4) Ln(volatility)	(5) F_trader	(6) Ln(M_HHI)
Intercept	-0.93***	-1.26***	1.18***	2.21***	-3.08***	0.81***
Gender	-0.01	-0.04***	0.06***	0.02	-0.05	0.04**
Age	0.01***	0.01***	-0.01***	-0.01***	0.01***	-0.01***
Level of education 1	0.11***	0.06***	-0.01*	-0.03	-0.01	-0.01
Level of education 2	0.01	0.07***	-0.08***	-0.17***	0.32***	-0.11***
Ln(PF value)	0.28***	0.29***	-0.20***	0.07***	0.11***	-0.16***
Trading experience	0.01***	0.01***	-0.01***	0.01***	0.01***	0.01***
Financial markets knowledge 1	0.05***	-0.02	-0.02	-0.01	0.23***	-0.10***
Financial markets knowledge 2	0.10***	-0.02*	-0.02	-0.04**	0.64***	-0.19***
Financial markets knowledge 3	0.14***	-0.1***	0.07***	-0.02	0.89***	-0.13***
Adjusted R^2	57.39%	52.23%	33.54%	5.74%	-	15.96%
Pseudo R^2	-	-	-		5.16%	-
N	20,285	20,285	20,285	20,285	20,285	20,285

The table reports the regression results for the relationship between our measures of diversification and subjective financial literacy in the A-test. The dependent variables of Regressions (1) to (6) are the natural logarithm of the number of different stocks traded during the sample period, the natural logarithm of the average number of stocks held in portfolio, the natural logarithm of the monthly average Herfindahl-Hirschman Index (computed as the sum of squared stock portfolio weights), the natural logarithm of the volatility of monthly portfolio returns, a binary variable set to 1 for investors who traded at least once investment fund shares, and the natural logarithm of the average modified Herfindahl-Hirschman Index (for which any position in funds is replaced by a portfolio of 50 equally-weighted securities). In the set of explanatory variables, 'Gender' is a dummy set to 1 for males, 'Age' is the investor's age in 2012, 'Level of education 1' is a dummy set to 1 for investors with a secondary/high school degree, 'Level of education 2' is a dummy set to 1 for investors with a university degree, 'Ln(PF value)' is the natural logarithm of the average monthly portfolio value (as a proxy of wealth) and 'Trading experience' is defined as the difference between the date of the last trade and the date of the first trade on stocks. The last three variables 'Financial markets knowledge 1', 'Financial markets knowledge 2', 'Financial markets knowledge 3' are dummies used respectively for the levels 1, 2 and 3 in the question about financial knowledge in the A-test as reported in Table 3. *** indicates significance at 1%; ** indicates significance at 5%; * indicates significance at 10%. 'N' gives the number of observations used in each model.

Table 11: Results for subjective financial literacy in the A-test & performance

	(1) Gross return	(2) Net return	(3) Gross Sharpe-Ratio	(4) Net Sharpe-Ratio	(5) Gross E-Sharpe-Ratio	(6) Net E-Sharpe-Ratio
Intercept	-1.53***	-2.65***	-0.09***	-0.16***	-0.10***	-0.17***
Gender	-0.28***	-0.39***	-0.01*	-0.02***	-0.03***	-0.03***
Age	0.02***	0.02***	0.01***	0.01***	0.01***	0.01***
Level of education 1	-0.17	-0.29**	-0.01	-0.02*	-0.02*	-0.02**
Level of education 2	-0.11	-0.10	0.01	-0.01	-0.01	-0.01
Ln(PF value)	0.11***	0.13***	0.01***	0.01***	0.01***	0.01***
Trading experience	0.01***	0.01***	0.01***	0.01***	0.01***	0.01***
Financial markets knowledge 1	0.12	0.17*	-0.01	-0.01	0.01	0.01
Financial markets knowledge 2	0.29***	0.27***	-0.01	-0.01	0.02***	0.01**
Financial markets knowledge 3	0.41***	0.29**	-0.01	-0.01	0.03***	0.02***
Adjusted R^2	1.92%	2.69%	0.71%	1.56%	1.82%	2.92%
Pseudo R^2	-	-	-	-	-	-
N	20,285	20,285	20,285	20,285	20,285	20,285

The table reports the regression results for the relationship between our measures of performance and subjective financial literacy in the A-test. The dependent variables of Regressions (1) to (6) are the geometric average of monthly gross returns, the geometric average of monthly net returns, the gross Sharpe ratio, the net Sharpe ratio, the gross excess Sharpe ratio, and the net excess Sharpe ratio. In the set of explanatory variables, 'Gender' is a dummy set to 1 for males, 'Age' is the investor's age in 2012, 'Level of education 1' is a dummy set to 1 for investors with a secondary/high school degree, 'Level of education 2' is a dummy set to 1 for investors with a university degree, 'Ln(PF value)' is the natural logarithm of the average monthly portfolio value (as a proxy of wealth) and 'Trading experience' is defined as the difference between the date of the last trade and the date of the first trade on stocks. The last three variables 'Financial markets knowledge 1', 'Financial markets knowledge 2', 'Financial markets knowledge 3' are dummies used respectively for the levels 1, 2 and 3 in the question about financial knowledge in the A-test as reported in Table 3. *** indicates significance at 1%; ** indicates significance at 5%; * indicates significance at 10%. 'N' gives the number of observations used in each model.

3.2.3 Discussion on subjective financial literacy and overconfidence

The literature often associates a high level of trading with overconfidence, i.e. people overtrade because they overestimate their actual skills and knowledge. The latter is usually considered as an investment mistake since it leads to poor net performance due to transaction costs. This view has been well summarized in Koestner et al. (2017) who define overtrading as one of the three investment mistakes the most cited in the literature and for which the significant negative effect on performance is well-documented. From that perspective, our results that show that investors who perceive themselves as highly literate tend to trade more looks totally consistent with overconfidence. However, the fact that those investors also display better performances, even after controlling for transaction costs and on a risk-adjusted basis, is much more surprising. To the best of our knowledge, this paper is the first to bring such empirical evidence. Dorn and Huberman (2005) and Graham et al. (2009) also find that investors with high subjective literacy tend to trade more. Nevertheless, these authors do not investigate the relationship between subjective financial literacy and trading performance. In this paper, we show that the higher trading activity of investors with higher levels of subjective financial literacy does turn into better performance.

The above new finding is consistent with a few papers that document a positive relationship between active trading and performance. For example, Shu et al. (2004) provide evidence that active trading is not necessarily detrimental to performance. These authors find a U-shape relationship between turnover and performance, suggesting that investors who trade the least/most earn higher returns than investors who exhibit an average trading activity. Such results are not consistent with the overconfidence hypothesis.

Distinguishing "smart" from "dumb" investors, Korniotis and Kumar (2013) provide consistent empirical evidence with our findings. In particular, these authors show that while both categories of investors display portfolio distortions, the underlying reasons differ: overtrading by "dumb" investors reflects behavioral biases, although the high trading activity of the "smart" ones reveals superior information. Korniotis and Kumar (2013) explain this finding by the strategy adopted by the"smart" investors. The latter combine both stock portfolio concentration and active trading to generate positive abnormal returns. The authors suggest that portfolio concentration eases "smart" investors' attention, information gathering as well as processing, which enables them to be better informed on a small set of stocks and trade more actively with profits. The investors who report higher levels of financial literacy in our sample seem to adopt a similar investment strategy since we find that those investors invest in investment funds and concentrate their stock portfolios on a small number of securities. This could enable them to benefit from risk diversification through funds and ease their information-gathering process to trade more actively on this small set of stocks.

4 Conclusion

Using survey data available in the MiFID tests, we investigate the relationship between subjective financial literacy and actual trading behavior. For that purpose, we analyse a sample of 20,285 retail investors who traded online during the 2003-2012 period and characterize their trading behaviors along three different aspects: experience and familiarity with financial markets, diversification and performance.

Regarding subjective financial literacy itself, our results provide evidence of overall consistency across investors' answers: investors who report a high literacy in one MiFID test are much more likely to do so in the other one. Retail investors are then consistent when self-reporting their financial literacy online. More importantly, we show that this piece of information provided by the investors themselves is helpful to characterize their actual trading behavior. Investors who report higher levels of financial literacy tend to invest *smarter*. Specifically, they trade more on both stocks and complex instruments, and they are less exposed to the disposition effect, which is consistent with higher experience. In addition, they tend to concentrate their stock portfolios on a small set of securities and achieve global diversification through investment funds holding. Finally, investors with a higher subjective financial literacy display higher both gross and net returns as well as higher excess Sharpe ratios. All these results hold even when we control for gender, age, portfolio value, trading experience, and education.

Our findings are not consistent with overconfidence because the higher trading activity of investors with higher levels of subjective financial literacy in our sample does result in better performance. This new empirical evidence is consistent with the strategy adopted by the "smart" investors in Korniotis and Kumar (2013). Those investors invest in investment funds and concentrate their stock portfolios on a small number of securities. This behavior could enable them to benefit from risk diversification through funds and ease their information-gathering process to trade more actively on this small set of stocks with profits.

This paper brings relevant insights for both policy making and understanding retail investors' behavior. Since subjective literacy reported in the MiFID tests is informative to characterize retail investors, it should deserve more attention in that perspective. Using subjective literacy could help investment firms provide the most suitable services to their retail clients. Generally speaking, this paper also opens new areas of research such as the role of opinions, perceptions and beliefs in the individuals' financial decision-making process.

5 Appendix

Table 12: Subjective financial literacy in the A-test vs. in the S-test (2)

A-test		S-test			
		0	1	2	Total
0	(#)	1,480	2,260	322	4,062
	(%)	7.30	11.14	1.59	20.02
	(r%)	36.44	55.64	7.93	
	(c%)	50.79	19.99	5.31	
1	(#)	858	4,146	698	5,702
	(%)	4.23	20.44	3.44	28.11
	(r%)	15.05	72.71	12.24	
	(c%)	29.44	36.67	11.51	
2	(#)	473	4,526	3,267	8,266
	(%)	2.33	22.31	16.11	40.75
	(r%)	5.72	54.75	38.52	
	(c%)	16.23	40.03	53.88	
3	(#)	103	375	1,777	2,255
	(%)	0.51	1.85	8.76	11.12
	(r%)	4.57	16.63	78.80	
	(c%)	3.53	3.32	29.30	
Total	(#)	2,914	11,307	6,064	20,285
	(%)	14.37	55.74	29.89	100.00

Statistic	Value	P-value
χ^2	6,185	<.0001
Spearman's rank correlation	0.49	<.0001

This contingency table reports respectively, for each pair of answers, the empirical frequencies (#), the total percentages (%), the row percentages (r%) and the column percentages (c%). Answers for the A-test are positioned in rows while those for the S-test are in columns. For the question in the A-test, levels 0 to 3 refer to the answers to the question about financial knowledge in the A-test as reported in Table 3. For the question in the S-test, levels 0 to 2 refer to the answers to the question about knowledge and experience about "complex" instruments in the S-test as reported in Table 3. The results for the Chi-Square test for the null hypothesis of independence are also provided as well as the results for the Spearman's rank correlation.

Table 13: Results for subjective literacy in the S-test & experience and familiarity with markets

	(1) Ln(total_trades)	(2) Ln(stock_trades)	(3) Ln(1+turnover)	(4) O_trader	(5) DE
Intercept	-0.03	-0.59***	2.52***	-3.81***	31.08***
Gender	0.10***	0.07***	0.10***	0.04	-1.43**
Age	0.01	-0.01	-0.01***	-0.01***	-0.17***
Level of education 1	0.26***	0.18***	0.12***	0.43***	-0.49
Level of education 2	0.09***	-0.02	-0.12***	0.48***	-1.93***
Ln(PF value)	0.30***	0.35***	0.05***	0.16***	-0.81***
Trading experience	0.02***	0.02***	-0.01***	0.01***	0.01
"Complex" instruments knowledge 1	0.03	-0.05**	-0.01	0.35***	-0.25
"Complex" instruments knowledge 2	0.31***	0.03	0.09***	1.36***	-1.90***
Adjusted R^2	43.88%	54.18%	2.52%	-	1.81%
Pseudo R^2	-	-	-	11.37%	-
N	20,285	20,285	20,285	20,285	20,285

The table reports the regression results for the relationship between our measures of experience and familiarity with financial markets and subjective financial literacy in the S-test. The dependent variables of Regressions (1) to (5) are the natural logarithm of the total number of trades across instruments, the natural logarithm of the number of stock trades, the natural logarithm of (1+turnover), a binary variable set to 1 when the investor traded at least once options or warrants, and the investor's DE measured as the difference between the proportion of gains realized and the proportion of losses realized. In the set of explanatory variables, 'Gender' is a dummy set to 1 for males, 'Age' is the investor's age in 2012, 'Level of education 1' is a dummy set to 1 for investors with a secondary/high school degree, 'Level of education 2' is a dummy set to 1 for investors with a university degree, 'Ln(PF value)' is the natural logarithm of the average monthly portfolio value (as a proxy of wealth) and 'Trading experience' is defined as the difference between the date of the last trade and the date of the first trade on stocks. The last two variables '"Complex" instruments knowledge 1', '"Complex" instruments knowledge 2' are dummies used respectively for the levels 1 and 2 in the question about knowledge and experience in complex instruments in the S-test as reported in Table 3. *** indicates significance at 1%; ** indicates significance at 5%; * indicates significance at 10%. 'N' gives the number of observations used in each model.

Table 14: Results for subjective financial literacy in the S-test & diversification

	(1) Ln(n_stocks)	(2) Ln(n_stocks_PF)	(3) Ln(HHI)	(4) Ln(volatility)	(5) F_trader	(6) Ln(M_HHI)
Intercept	-0.95***	-1.24***	1.18***	2.23***	-3.32	0.87***
Gender	-0.01	-0.04***	0.05***	0.02	-0.04	0.04*
Age	0.01***	0.01***	-0.01***	-0.01***	0.01***	-0.01***
Level of education 1	0.14***	0.05***	-0.04**	-0.04	0.21***	-0.07***
Level of education 2	0.05***	0.06***	-0.09***	-0.18***	0.57***	-0.19***
Ln(PF value)	0.28***	0.29***	0.20***	0.07***	0.11***	-0.15***
Trading experience	0.01***	0.01***	-0.01***	0.01***	0.01***	0.01***
"Complex" instruments knowledge 1	0.01	-0.04***	0.01	-0.03	0.28***	-0.07***
"Complex" instruments knowledge 2	0.07***	-0.08***	0.05***	-0.05**	0.76***	-0.12***
Adjusted R^2	57.33%	52.23%	33.46%	5.74%	-	15.62%
Pseudo R^2	-	-	-		4.99%	-
N	20,285	20,285	20,285	20,285	20,285	20,285

The table reports the regression results for the relationship between our measures of diversification and subjective financial literacy in the S-test. The dependent variables of Regressions (1) to (6) are the natural logarithm of the number of different stocks traded during the sample period, the natural logarithm of the average number of stocks held in portfolio, the natural logarithm of the monthly average Herfindahl-Hirschman Index (computed as the sum of squared stock portfolio weights), the natural logarithm of the volatility of monthly portfolio returns, a binary variable set to 1 for investors who traded at least once investment fund shares, and the natural logarithm of the average modified Herfindahl-Hirschman Index (for which any position in funds is replaced by a portfolio of 50 equally-weighted securities). In the set of explanatory variables, 'Gender' is a dummy set to 1 for males, 'Age' is the investor's age in 2012, 'Level of education 1' is a dummy set to 1 for investors with a secondary/high school degree, 'Level of education 2' is a dummy set to 1 for investors with a university degree, 'Ln(PF value)' is the natural logarithm of the average monthly portfolio value (as a proxy of wealth) and 'Trading experience' is defined as the difference between the date of the last trade and the date of the first trade on stocks. The last two variables '"Complex" instruments knowledge 1', '"Complex" instruments knowledge 2' are dummies used respectively for the levels 1 and 2 in the question about knowledge and experience in complex instruments in the S-test as reported in Table 3. *** indicates significance at 1%; ** indicates significance at 5%; * indicates significance at 10%. 'N' gives the number of observations used in each model.

Table 15: Results for subjective financial literacy in the S-test & performance

	(1) Gross return	(2) Net return	(3) Gross Sharpe-Ratio	(4) Net Sharpe-Ratio	(5) Gross E-Sharpe-Ratio	(6) Net E-Sharpe-Ratio
Intercept	-1.59***	-2.7***	-0.08***	-0.15***	-0.09***	-0.17***
Gender	-0.28***	-0.38***	-0.01*	-0.02***	-0.03***	-0.03***
Age	0.02***	0.01***	0.01***	0.01***	0.01***	0.01***
Level of education 1	-0.07	-0.19	-0.01	-0.02**	-0.01	-0.02**
Level of education 2	-0.01	0.01	-0.01	-0.01	0.01	0.01
Ln(PF value)	0.11***	0.13***	0.01***	0.01***	0.01***	0.01***
Trading experience	0.01***	0.01***	0.01***	0.01***	0.01***	0.01***
"Complex" instruments knowledge 1	0.04	0.01	-0.01	-0.01	-0.01	-0.01
"Complex" instruments knowledge 2	0.28***	0.19**	-0.01	-0.01	0.01	-0.01
Adjusted R^2	1.91%	2.68%	0.72%	1.57%	1.77%	2.89%
Pseudo R^2	-	-	-	-	-	-
N	20,285	20,285	20,285	20,285	20,285	20,285

The table reports the regression results for the relationship between our measures of performance and subjective financial literacy in the S-test. The dependent variables of Regressions (1) to (6) are the geometric average of monthly gross returns, the geometric average of monthly net returns, the gross Sharpe ratio, the net Sharpe ratio, the gross excess Sharpe ratio, and the net excess Sharpe ratio. In the set of explanatory variables, 'Gender' is a dummy set to 1 for males, 'Age' is the investor's age in 2012, 'Level of education 1' is a dummy set to 1 for investors with a secondary/high school degree, 'Level of education 2' is a dummy set to 1 for investors with a university degree, 'Ln(PF value)' is the natural logarithm of the average monthly portfolio value (as a proxy of wealth) and 'Trading experience' is defined as the difference between the date of the last trade and the date of the first trade on stocks. The last two variables '"Complex" instruments knowledge 1', '"Complex" instruments knowledge 2' are dummies used respectively for the levels 1 and 2 in the question about knowledge and experience in complex instruments in the S-test as reported in Table 3. *** indicates significance at 1%; ** indicates significance at 5%; * indicates significance at 10%. 'N' gives the number of observations used in each model.

References

Abreu, M., Mendes, V., 2012. Information, overconfidence and trading: Do the sources of information matter? Journal of Economic Psychology 33 (4), 868–881.

Allgood, S., Walstad, W., 2013. Financial literacy and credit card behaviors: A cross-sectional analysis by age. Numeracy 6 (2), Article 3.

AMF, 2010. Evaluation of MiFID questionnaires in France. Study for the Autorité des Marchés Financiers.

Barber, B. M., Odean, T., 2000. Trading is hazardous to your wealth: The common stock investment performance of individual investors. Journal of Finance 55 (2), 773–806.

Barber, B. M., Odean, T., 2001. Boys will be boys: Gender, overconfidence, and common stock investment. Quarterly Journal of Economics 116 (1), 261–292.

Barber, B. M., Odean, T., 2013. The behavior of individual investors. In: Handbook of the Economics of Finance, volume 2, part B, edited by George M. Constantinides, Milton Harris and Rene M. Stulz, 1533-1570.

Bertrand, M., Mullainathan, S., 2001. Do people mean what they say? Implications for subjective survey data. American Economic Review 91 (2), 67–72.

Blume, M. E., Friend, I., 1975. The asset structure of individual portfolios and some implications for utility functions. Journal of Finance 30 (2), 585–603.

Boolell-Gunesh, S., Broihanne, M., Merli, M., 2012. Sophistication of individual investors and disposition effect dynamics. Finance 33 (1), 9–37.

Bruine de Bruin, W., 2011. Framing effects in surveys: How respondents make sense of the questions we ask. In: Perspectives on Framing, edited by Gideon Keren, 303-325. New-York, NY. Taylor and Francis Group, Psychology Press.

Bucher-Koenen, T., Lusardi, A., Alessie, R., van Rooij, M., 2012. How financially literate are women? Some new perspectives on the gender gap. Network for Studies on Pensions, Aging and Retirement. Panel Paper 31.

Campbell, J. Y., 2006. Household finance. Journal of Finance 61 (4), 1553–1604.

Christelis, D., Jappelli, T., Padula, M., 2010. Cognitive abilities and portfolio choice. European Economic Review 54 (1), 18–38.

Dhar, R., Zhu, N., 2006. Up close and personal: An individual level analysis of the disposition effect. Management Science 52 (5), 726–740.

Dorn, D., Huberman, G., 2005. Talk and action: What individual investors say and what they do. Review of Finance 9 (4), 437–481.

ESMA, 2012. Guidelines on certain aspects of the MiFID suitability requirements. Final report.

Feng, L., Seasholes, M. S., 2005. Do investor sophistication and trading experience eliminate behavioral biases in financial markets? Review of Finance 9 (3), 305–351.

FSA, 2011. Assessing suitability: Establishing the risk a customer is willing and able to take and making a suitable investment selection. Finalised Guidance 11/5.

FSMA, 2014. Devoir de diligence constatations et positions dégagées à l'issue de plusieurs inspections. Rapport aout 2014.

Glaser, M., 2003. Online broker investors: Demographic information, investment strategy, portfolio positions, and trading activity. Available at SSRN 975985.

Glaser, M., Weber, M., 2007. Overconfidence and trading volume. Geneva Risk and Insurance Review 32 (1), 1–36.

Glaser, M., Weber, M., 2009. Which past returns affect trading volume? Journal of Financial Markets 12 (1), 1–31.

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

Graham, J. R., Harvey, C. R., Huang, H., 2009. Investor competence, trading frequency, and home bias. Management Science 55 (7), 1094–1106.

Guiso, L., Jappelli, T., 2008. Financial literacy and portfolio diversification. Working Paper, European University Institute.

Hackethal, A., Haliassos, M., Jappelli, T., 2012. Financial advisors: A case of babysitters? Journal of Banking and Finance 36 (2), 509–524.

Haliassos, M., Bertaut, C. C., 1995. Why do so few hold stocks? Economic Journal 105 (432), 1110–29.

Hilgert, M. A., Hogarth, J. M., Beverly, S. G., 2003. Household financial management: The connection between knowledge and behavior. Federal Reserve Bulletin 89 (7), 309–322.

Hoffmann, A. O., Post, T., Pennings, J. M., 2013. Individual investor perceptions and behavior during the financial crisis. Journal of Banking & Finance 37 (1), 60–74.

Hsiao, Y.-J., Tsai, W.-C., 2018. Financial literacy and participation in the derivatives markets. Journal of Banking & Finance 88, 15–29.

Kimball, M., Shumway, T., 2006. Investor sophistication, and the participation, home bias, diversification, and employer stock puzzles. Ann Arbor: Unpublished Manuscript, University of Michigan.

Koestner, M., Loos, B., Meyer, S., Hackethal, A., 2017. Do individual investors learn from their mistakes? Journal of Business Economics 87 (5), 669–703.

Korniotis, G. M., Kumar, A., 2013. Do portfolio distortions reflect superior information or psychological biases? Journal of Financial and Quantitative Analysis 48 (1), 1–45.

Kumar, A., Lee, C., 2006. Retail investors sentiment and returns comovements. Journal of Finance 61 (5), 2451–2486.

Linciano, N., Soccorso, P., 2012. Assessing investors' risk tolerance through a questionnaire. CONSOB Discussion Papers.

Lusardi, A., 2011. Americans' financial capability. National Bureau of Economic Research. Working Paper 17103.

Lusardi, A., Mitchell, O. S., 2007. Baby boomer retirement security: The roles of planning, financial literacy, and housing wealth. Journal of Monetary Economics 54 (1), 205–224.

Lusardi, A., Mitchell, O. S., 2008. Planning and financial literacy: How do women fare? American Economic Review 98 (2), 413–417.

Lusardi, A., Mitchell, O. S., 2011a. Financial literacy and planning: Implications for retirement well-being. In: Financial literacy: Implications for retirement security and the financial marketplace, edited by Olivia S. Mitchell and Annamaria Lusardi, 17-39. Oxford and New-York: Oxford University Press.

Lusardi, A., Mitchell, O. S., 2011b. Financial literacy around the world: An overview. Journal of Pension Economics and Finance 10 (4), 497–508.

Lusardi, A., Mitchell, O. S., 2014. The economic importance of financial literacy: Theory and evidence. Journal of Economic Literature 52 (1), 5–44.

Marinelli, N., Mazzoli, C., 2011. An insight into suitability practice: Is a standard questionnaire the answer? In: Bank Strategy, Governance and Ratings, edited by Philip Molyneux, 217-245. Palgrave Macmillan.

Mazzoli, C., Marinelli, N., 2014. Determinants of risk-suitable investment portfolios: Evidence from a sample of italian householders. International Journal of Finance & Banking 1 (1), 13–28.

Merkle, C., Weber, M., 2014. Do investors put their money where their mouth is? Stock market expectations and investing behavior. Journal of Banking & Finance 46, 372–386.

Odean, T., 1998. Are investors reluctant to realize their losses? Journal of finance 53 (5), 1775–1798.

Sekita, S., 2011. Financial literacy and retirement planning in japan. Journal of Pension Economics and Finance 10 (4), 637–656.

Shefrin, H., Statman, M., 1984. The disposition to sell winners too early and ride losers too long: Theory and evidence. Journal of Finance 40 (3), 777–790.

Shu, P.-G., Chiu, S.-B., Chen, H.-C., Yeh, Y.-H., 2004. Does trading improve individual investor performance? Review of Quantitative Finance and Accounting 22 (3), 199–217.

Tversky, A., Kahneman, D., 1981. The framing of decisions and the psychology of choice. Science 211 (4481), 453–458.

Van Rooij, M., Lusardi, A., Alessie, R., 2011. Financial literacy and stock market participation. Journal of Financial Economics 101 (2), 449–472.

Vissing-Jorgensen, A., 2003. Perspectives on behavioral finance: Does "irrationality" disappear with wealth? Evidence from expectations and actions. In: NBER Macroeconomics Annual 2003, edited by M. Gertler and K. Rogoff, 139-208. The MIT Press.

Volpe, R. P., Kotel, J. E., Chen, H., 2002. A survey of investment literacy among online investors. Financial Counseling and Planning 13 (1), 1–13.

Xia, T., Wang, Z., Li, K., 2014. Financial literacy overconfidence and stock market participation. Social Indicators Research 119 (3), 1–13.

Appetite for Information in Mandatory Profiling of Individual Investors

Anthony Bellofatto[a] and Marie-Hélène Broihanne[b]

This version: July 20, 2018

Abstract

Come into force in November 2007, MiFID questionnaires offer a natural field experiment to analyze the relationships between retail investors' attitude towards financial information and their trading activity. Under a random matching procedure that controls for socio-demographics, financial experience, education and various MiFID answers, we analyze the trading characteristics of investors who only differ from others on the side of their "appetite for information". We conjecture that the investors who voluntary ask for financial information have revealed de facto a particular feature that may be indicative of their trading behavior. Our results show that the investors who display an appetite for information tend to invest smarter than their counterparts. Actually, they execute less daytrades, are less subject to the Disposition Effect, hold better diversified portfolios, and are more active on "complex" instruments. They *in fine* earn higher (risk-adjusted) returns.

JEL Classification: D14, D83, G11, G28, G40

Keywords: retail investors, information acquisition, financial knowledge, MiFID

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. Any errors are the full responsibility of the authors. The authors wish to thank C. D'Hondt for her useful comments and her precious work on the database as well as M. Merli and P. Roger for their comments. They also wish to thank Ryan L. Davis for his useful discussion of the paper as well as the participants at the 66th Annual Meeting of the Midwest Finance Association, at the 2017 Joint Conference of the Academy of Entrepreneurial Finance and the Academy of Behavioral Finance & Economics and at the 34th International Conference of the French Finance Association. Comments are welcome.

(a)Louvain Finance (IMMAQ), Louvain School of Management, Université catholique de Louvain - Address: Chaussée de Binche 151, 7000 Mons, Belgium - Email: anthony.bellofatto@uclouvain.be;

(b)LaRGE Research Center, EM Strasbourg Business School, University of Strasbourg - Address: Avenue de la Forêt Noire 61, 67085 Strasbourg Cedex, France - Email: mhb@unistra.fr (corresponding author).

1 Introduction

In November 2007, the Markets in Financial Instruments Directive (MiFID) came into force.[1] Its goal is to increase the level of protection of retail investors by requiring investment firms to deliver the most suitable services to their clients. In this perspective, investment firms operating in the European Union are now obliged to collect information about their retail clients through the so-called "MiFID tests". While the directive requires investment firms to gather relevant information about their clients' profile, the quantity and the nature of the information to be collected depend on the service asked by the retail investor. As illustrated in Figure 1, the Directive determines three types of services (Committee of European Securities Regulators (2008)): Execution of orders, financial advice and portfolio management.

The investors who ask the banks to execute transactions on "complex" instruments have to fulfill the *Appropriateness test* (hereafter the A-test) that ensures that the customer has the necessary experience and knowledge to understand the risks involved in "complex" financial instruments before investing. However, the investors who also ask for financial advice or portfolio management have in addition to fulfill the *Suitability test* (hereafter the S-test). Assessment of suitability involves ensuring that the instruments and services offered meet the investor's objectives, financial capacity as well as his knowledge and experience in financial instruments. Henceforth, MiFID requirements offer a natural field to investigate the relationships between a purposeful need for information, i.e. asking for advice, and the trading behavior of retail investors.

Our paper addresses this topic of research by using a database coming from an online Belgian brokerage house including the MiFID questionnaires records of 14,155 retail investors and their trading activity over the 2008-2012 period. Since our data are provided by an online brokerage house that does not offer any portfolio management service during the sample period, the investors in our sample have either fulfilled the A-test (hereafter A-investors) to execute transactions or fulfilled the A-test and S-test (hereafter S-investors) to have access to financial advice. As for the online brokerage house under study, the S-investors get an access to an information tool on stocks as financial advice. Bearing the supplementary cost of time needed to fulfill the S-test,[2] the S-investors have shown a willingness to access a service higher than orders execution alone ("premium service"). By doing so, we conjecture that those investors

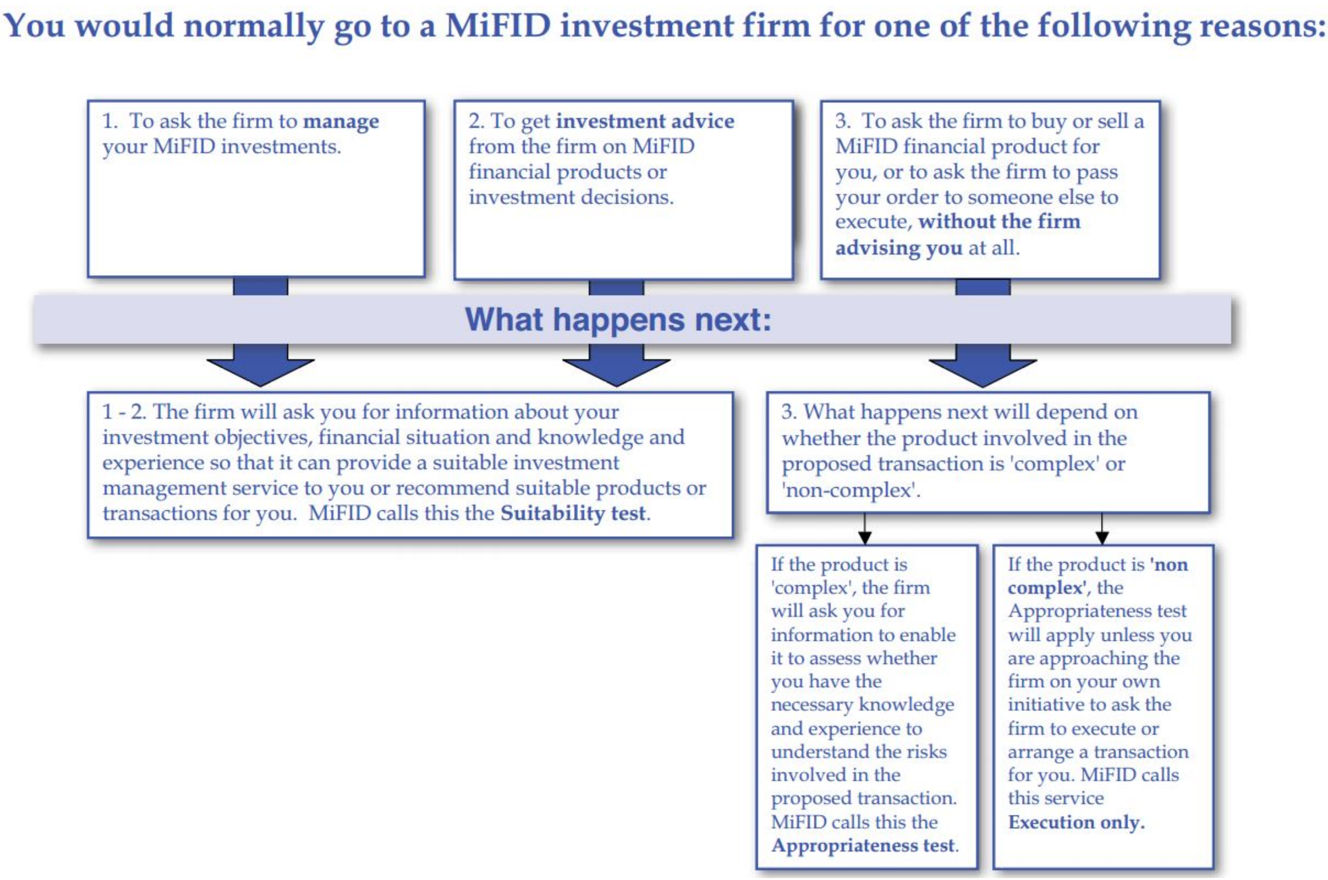
[1]MiFID is applied in member states of the European Economic area (28 European member states and Iceland, Norway and Liechtenstein). From January 2018, MiFID I (2004/39/EC) has been replaced by MiFID II (2014/65/UE) which carries similar objectives.

[2]It took an average of 30 minutes to fulfill the S-test.

have revealed a distinct feature about their personality, that we call “appetite for information”, that may be indicative of a specific trading behavior. On the contrary, the A-investors have neglected a free access to professional recommendations. The aim of our paper is therefore to investigate whether the appetite for information revealed (or not) by a retail investor when fulfilling the MiFID questionnaires is related to his trading behavior.

The information content of MiFID questionnaires answers has already been shown for risk profile analysis (Mazzoli and Marinelli (2014)), sentiment trading (D’Hondt and Roger (2017)), mental accounting (Broihanne and Orküt (2018)) and subjective financial literacy (Bellofatto et al. (2018)). However, this paper is the first to investigate whether investors’ behavior is consistent with their attitude towards financial information revealed through MiFID questionnaires.

Figure 1: MiFID services



The figure exhibits the three types of service recognized by the MiFID Directive and information investors need to provide accordingly. Source: Committee of European Securities Regulators (2008)

Our paper is part of the literature on financial information. This strand of literature is characterized by a huge debate concerning the relationship between financial knowledge needs and financial knowledge acquisition. Although some authors argue that only financial education is needed, others suggest that financial advice could be a good solution to the lack of financial knowledge among individual investors (a.o. Bucher-Koenen and Koenen (2010) and Georgarakos and Inderst (2014)). Financial advice and financial knowledge would therefore be substitutes. In the opposite vein, another approach involves considering financial advice and financial knowledge as complementary (Calcagno and Monticone (2015)). However, our paper differs from those papers since we do not investigate the benefits of financial information (received from a financial advisor) *per se* but the effect of the personal characteristics that lead individual investors to voluntary ask (or not) for useful financial information on their trading behavior.

Our results indicate that a purposeful need for information (or not) is indicative of trading behavior. We find that the investors who reveal an appetite for financial information invest smarter. They execute less daytrades, are less subject to the Disposition Effect, are more active on "complex" instruments, which indicates a higher level of financial experience and familiarity with financial markets, and hold better diversified portfolios. At the opposite, the A-investors display a more "intuitive" trading behavior. They concentrate their trades on a lower number of stocks, execute more daytrades and roundtrips on this stock set and are less attracted by "other-than-stocks" instruments. This "intuitive" trading behavior is particularly in line with D'Hondt and Roger (2017) who show that the A-investors are more prone to sentiment trading than the S-investors and that their trading behavior are more predictive of future returns on a long-short portfolio based on size. Evidence of this trading behavior difference may explain why the S-investors earn significantly higher (risk-adjusted) returns. These findings hold even under a random matching procedure that controls for socio-demographic data, financial experience, education and various survey answers.

With these results at hand, we contribute to the literature on the relationships between information acquisition and trading activity. While several papers (a.o Abreu and Mendes (2012) and Tauni et al. (2015)) report empirical[3] evidence of the positive relationship between information acquisition and trading activity, they investigate this relationship only descriptively

[3]This topic of research has already been addressed from a theoretical point of view. In theoretical models, information investment is rational for investors as long as the cost of searching information exceeds its marginal benefit. According to authors (a.o. Grossman and Stiglitz (1980), Karpoff (1986) and Holthausen and Verrecchia (1990)), the investors who spend more time searching for information tend to compensate the cost of information by taking more risky positions and by trading more. However, Argentesi et al. (2010) have a slightly different perspective. They argue that: "*The fact that more information is collected by investors does not necessarily*

since they do not have trading records but only written answers to surveys.[4] Furthermore, they only focus on trading frequency and dot not analyze trading behavior in a broader sense or even trading performance. To the best of our knowledge, the paper of Guiso and Jappelli (2006) is the only exception. In contrast to Abreu and Mendes (2012) and Tauni et al. (2015), they use a detailed database including the trading records and survey answers of 1,834 customers of a leading Italian commercial bank to investigate the determinants and the effect of information acquisition on trading behavior and performance. They provide evidence of a negative relationship between information acquisition and returns, supporting the overconfidence hypothesis. Looking at investors' behavior, Guiso and Jappelli (2006) find that the investors who spend more time searching for financial information are associated to frequent trading, less diversified portfolios and lower tendency to delegate.

Besides the fact that we investigate a larger and more recent sample, the difference with the above paper lies in the proxies used to measure investors' attitude towards information acquisition. While Guiso and Jappelli (2006) measure through a questionnaire the time spent for acquiring financial news whatever the source of information (reading the newspapers, surfing on the web, ...), the S-investors in our sample voluntary fulfill the S-test to have an access to a directly usable information tool. Since these investors have produced an "effort" to access an information tool, our measure may be more indicative of their appetite for information.

The remainder of this paper is structured as follows. Section 2 describes the data. Section 3 describes the methodology we use. We report our empirical work and its results in Section 4. Section 5 concludes.

2 Data and Sample

The database is provided by an online Belgian brokerage house and encompasses the trading activity of 14,155 retail investors over the January 2008 - March 2012 period.[5] Two datasets composed the data. The first one contains information about the investors, that we classify into

imply that more trading will follow (for instance, because information may just suggest that it is optimal not to trade)".

[4] Abreu and Mendes (2012) use a survey conducted in 2000 by the Portuguese Securities Market Commission in which 1,559 investors were interviewed. Tauni et al. (2015) analyze the survey results of 333 individual investors in Chinese future markets. Both papers ask a question like "How often do you buy and sell financial assets?".

[5] D'Hondt and Roger (2017) and Bellofatto et al. (2018) analyze the same database while they focus on a different sample in accordance with their research question.

three categories. The first category includes socio-demographic data: year of birth, gender and spoken language. The second category encompasses the answers to the A-test while the third category contains the answers to the S-test. The second dataset is made of detailed information about the investors' trading activity on stocks, funds, options, warrants, and bonds. For the purpose of our study, we use information about the stock trading activity to build end-of-month portfolios for each investor in the sample.[6] We complement this dataset with Eurofidai[7] and Bloomberg[8] historical data to compute the market value of the end-of-month portfolios.

2.1 Trading activity

Our sample of investors has made 654,678 trades on 5,959 different stocks,[9] which represents about 154,000 trades in a typical year and about 13,000 trades in a typical month. The investors in our sample are net buyers since 60% of the trades are purchases and 40% are sales.

Table 1 presents descriptive statistics for trading activity. The average investor completes 44 trades on 12 different stocks over a 25 months trading period.[10] The typical investor makes about 1.4 times daytrading[11] and trades on average 3.37 times the same stock over his whole period of trading. As for the trading activity on "complex" instruments, the average investor completes about 7 trades on investment fund shares, 8 trades on options or warrants and almost no trade on bonds. All the above variables are positively skewed since the means are substantially larger than the medians.

[6] In our sample, we only consider investors who began to trade at the brokerage house under study after January the 1st, 2008, i.e. after the MiFID tests came into force. We therefore assume that those investors first fulfilled the tests before trading.

[7] www.Eurofidai.org

[8] www.Bloomberg.com

[9] We focus on stocks for which an ISIN code is available. For stocks traded in foreign currencies, we use exchange rates to convert monetary volumes into euros.

[10] We compute the trading experience as the difference between the last trade date and the first trade date available in the sample. As in Glaser and Weber (2009) we exclude from our sample investors with less than 5 months of trading activity.

[11] We compute 1 daytrade each time an investor makes a purchase and a sale on the same stock on the same day.

Table 1: Descriptive statistics for trading activity

	Mean	Median	Q1	Q3
Number of stock trades	44	18	8	45
Number of different stocks traded	12	7	4	15
Trading experience (in months)	25	24	14	35
Number of daytrades	1.43	0	0	0
Average number of trades on the same stock	3.37	2.4	1.75	3.64
Number of fund trades	7.04	0	0	0
Number of option trades	8.31	0	0	0
Number of bond trades	0.08	0	0	0

The table reports the cross-sectional mean, median, lower and upper quartiles for trading activity variables on a per investor basis over the sample period. 'Number of stock trades' is the number of trades executed on stocks. 'Number of different stocks traded' is the number of different stocks traded during the whole trading period. 'Trading experience' is computed as the difference between the last trade date and the first trade date available in the sample. It is expressed in number of months. 'Number of daytrades' is the number of times an investor makes a purchase and a sale on the same stock on the same day. 'Average number of trades on the same stock' is the average number of trades an investor makes on the same stock. 'Number of fund trades' is the number of trades executed on investment fund shares. 'Number of option trades' is the number of trades executed on both options and warrants. 'Number of bond trades' is the number of trades executed on bonds.

Table 2 shows statistics computed on binary variables. While 21.79% of the investors trade investment fund shares, 18.26% of them trade options or warrants, but only 3.16% of them trade bonds.

Table 2: Frequencies of trading activity on "complex" instruments

	0	1
Funds_trader	78.21%	21.79%
Options_trader	81.74%	18.26%
Bonds_trader	96.84%	3.16%

The table reports frequencies of trading activity on "complex" instruments over the sample period. 'Funds_trader' is set to 1 when the investor made at least one trade on investment fund shares. 'Options_trader' is set to 1 when the investor made at least one trade on either options or warrants. 'Bonds_trader' is set to 1 when the investor made at least one trade on bonds.

We use data on the trading activity on stocks to build end-of-month portfolios. Combining our data with historical market data, we compute the monthly average number of stocks held in portfolio, the monthly average portfolio value as well as the monthly returns.[12]

Table 3 reports descriptive statistics for the above measures. We know that the average investor holds a four-stock portfolio, this underdiversification being in line with Kumar and Lee (2006) and Polkovnichenko (2010) for the US and Broihanne et al. (2016) in Europe (France). The median of 2.76 is also consistent with Goetzmann and Kumar (2008) who find that more than 50% of the retail investors in their sample hold only one to three stocks. The average end-of-month portfolio value is about 22,000 euros with a median of 6,500 euros. As for the variables reported in Table 1, all these portfolio-based variables are positively skewed.

The average investor earns a monthly gross (net[13]) return of 0.40 (-0.40)% in a typical month.[14] The average monthly volatility of the returns is about 18% with a median of 11.22%. The mean and median volatility are larger than the ones reported in Dorn and Huberman (2005) but the specificity of our sample period may explain the difference.[15] In addition, the mean value of 18% is slightly higher than the monthly realized volatility of the BEL20 and CAC40 indices, which may represent appropriate benchmarks for Belgian investors, over the 2008-2012 period.

The descriptive statistics depict a huge heterogeneity in the behavior and performance of the investors in our sample, which is particularly in line with the extant literature on retail investors (Barber and Odean (2013)).

[12]To compute trading performance, we make one assumption commonly used in the literature (Barber and Odean (2000), Barber and Odean (2001a), Shu et al. (2004) and Glaser and Weber (2007)): we assume that all transactions take place on the last day of the month. Barber and Odean (2000) have shown that this simplifying assumption do not bias the measurement of portfolio performance.

[13]Only explicit transactions costs are taken into account.

[14]For each investor, we compute a geometric average return.

[15]They investigate the 1995-2000 period while we analyze the post-2008 period.

Table 3: Descriptive statistics for end-of-month portfolio data

	Mean	Median	Q1	Q3
Number of stocks	4.25	2.76	1.36	5.29
Portfolio value (€)	22,005	6,490	2,195	17,779
Gross return (%)	0.40	0.23	-1.47	1.98
Net return (%)	-0.40	-0.22	-2.21	1.48
Volatility (%)	18.01	11.22	7.17	18.29

The table reports the cross-sectional mean, median, lower and upper quartiles for portfolio data variables on a per investor basis over the sample period. 'Number of stocks' is the monthly average number of stocks held in portfolio. 'Portfolio value' is the monthly average end-of-month portfolio market value. 'Gross return' is the geometric average of the monthly gross returns. 'Net return' is the geometric average of the monthly net returns. 'Volatility' is the standard deviation of the monthly returns.

2.2 A- and S-investors

Our sample is composed of two categories of investors that we can distinguish on the services they have asked and on the information they have provided accordingly. On the one hand, 6,913 investors have asked the bank to only execute their transactions. These investors have, accordingly to the MiFID Directive, only fulfilled the Appropriateness test (A-investors). On the other hand, 7,242 investors have asked, in addition, to have an access to an information tool on stocks. Therefore they have also fulfilled the Suitability test (S-investors). As a consequence, while both groups execute trades by themselves, the S-investors have a free access to an investment advice tool on stocks. The only "cost" endured to access the information tool is bearing time (about 30 minutes) to fill in the S-test. In our case, the S-test under scrutiny is made of 11 questions and covers, in accordance with the MiFID Directive, the investor's objectives, financial capacity as well as his knowledge and experience in financial instruments.

Since both groups have fulfilled the A-test, information contained in this test can be used to characterize the overall profile of our investors. The A-test in our database consists of a list of categorical questions for which the investors have to select an answer.[16] In addition, we

[16] We have grouped some questions when they were related to the same topic. For example, the A-test contains detailed questions regarding the investor's knowledge of options, futures, warrants, structured products,... Due to low frequencies, we have decided to group them and to create a question related to the knowledge of "complex" instruments.

have information about some socio-demographic variables. Table 4 reports for each question the dispersion of the investors between the different categories.

From Table 4, we know that 8.04% of the investors have chosen the highest level when they had to estimate their level of financial markets knowledge. As for the second question, about 5% of the investors have evaluated their experience in options, structured products, forex and futures as “good”, while 9.98% as “average”, and no experience for the remaining investors. Furthermore, about 34% of the investors have already invested at least once in an option, a structured product, a forex instrument or a future. As for the level of education, 72% have stated to have a university degree or equivalent, while 21% a secondary/high school and 6% no degree. A majority of investors are males and speak Dutch as their main language. In addition, not reported in the table, the average investor is 44 years old. Finally, in our sample, about 17% of the investors have stated to have executive responsibilities.

Table 4: Statistics for investors' characteristics

	Empirical frequencies (%)
Self-estimated knowledge of financial markets	
Level 0	29.21
Level 1	30.99
Level 2	31.76
Level 3	8.04
Self-evaluated experience in "complex" instruments	
Level 0	84.71
Level 1	9.98
Level 2	5.31
Investment in "complex" instruments	
No	66.13
Yes	33.87
Level of education	
Level 0	6.09
Level 1	21.49
Level 2	72.42
Gender	
Female	14.80
Male	85.20
Language	
French-speaker	45.35
Dutch-speaker	50.77
English-speaker	3.88
Professional status	
Executive	16.67
Other	83.33
N	14,155

The table reports empirical frequencies for investors' characteristics. As for the self-estimated knowledge of financial markets, level 0 is associated with a basic knowledge while level 3 refers to an experienced investor who manages any aspect of financial markets. As for the self-evaluated experience in "complex" instruments, level 0, level 1 and level 2 corresponds respectively to 'no experience', an 'average experience' and a 'good experience' in options, structured products, forex and futures. Investment in "complex" instruments is 'yes' if the investor states to have already invested in options, structured products, forex or futures contracts, and 'no' otherwise. As for the level of education, level 0 corresponds to 'no degree', level 1 to a 'secondary school/sigh school degree' and level 3 to a 'university degree'. Gender is 'female' if the investor is a woman and 'male' if the investor is a man. Language is 'French-speaker' if the investor is a French-speaker as his main language, 'Dutch-speaker' if the investor is a Dutch-speaker as his main language and 'English-speaker' if the investor is an English-speaker as his main language. Professional status is 'executive' if the investor claims executive responsibilities and 'other' otherwise.

3 Methodology

The A- and S-investors execute trades by themselves but the S-investors have, in addition, voluntary asked to have an access to additional information. Asides from having fulfilled the A-test, the S-investors have endured the cost of filling in the S-test to have an access to the advice tool. Therefore, we conjecture that they have revealed a particular appetite for information that may be related to their trading behavior. On the contrary, the A-investors have neglected a free access to more financial information. The aim of this paper is therefore to compare the trading behavior of the A- and S-investors.

However, since the investors who ask for financial information may differ from the others on a large set of covariates, comparing the trading behavior of the A- and S-investors to study the "appetite for information effect" may be subject to the "omitted variable" bias.[17] As a consequence, a difference in the trading behavior between both groups of investors could be due to other investor-immanent effects that are correlated with the appetite for information.

In this perspective, Table 5 compares the A- and S-investors on the variables displayed in Table 4. Table 5 reports for each categorical variable the proportion by group of investors, the difference between both proportions as well as the significance of the difference.

Table 5 clearly suggests that the A- and S-investors significantly differ on the vast majority of the variables. We therefore need to control the effect of these variables to investigate the impact of the appetite for information. As stated in Stuart (2010), when estimating causal effects using observational data, it is desirable to replicate experiments as closely as possible by obtaining treated and control groups with similar covariates distributions. In this perspective, we apply a random matching method since it is frequently done in the literature (a.o. Gerhardt and Hackethal (2009) and Kramer (2012)).

According to Stuart (2010), the nearest-neighbor matching is one of the most common and easiest to implement matching method. In its simplest form, 1:1 nearest neighbor matching selects for each treated individual i the control individual with the smallest distance from individual i. The distance between two individuals is based on their respective propensity score that Rosenbaum and Rubin (1983) define as the probability to receive the treatment given the observed covariates. The propensity score has the advantage to facilitate the construction of

[17]On the paper of Bluethgen et al. (2008) that compares the trading behavior of advised and non-advised investors, Gerhardt and Hackethal (2009) claim that many aspects of the difference between advised and non-advised investors can be attributed to differences in investors' characteristics.

Table 5: Comparison of investors' characteristics between A- and S-investors

	A-investors (%)	S-investors (%)	Difference (%)
Self-estimated knowledge of financial markets			
Level 0	29.30	29.12	-0.18
Level 1	31.01	30.97	-0.04
Level 2	30.72	32.74	2.02***
Level 3	8.97	7.17	-1.80***
Self-evaluated experience in "complex" instruments			
Level 0	82.77	86.57	3.8***
Level 1	11.10	8.91	-2.19***
Level 2	6.13	4.52	-1.61***
Investment in "complex" instruments			
No	67.08	65.23	-1.85**
Yes	32.92	34.77	1.85**
Level of education			
Level 0	7.03	5.19	-1.84***
Level 1	22.90	20.15	-2.75***
Level 2	70.07	74.66	4.59***
Gender			
Female	18.91	10.88	-8.03***
Male	81.09	89.12	8.03***
Language			
French-speaker	47.62	43.19	-4.43***
Dutch-speaker	48.36	53.08	4.72***
English-speaker	4.02	3.73	-0.29
Professional status			
Executive	15.15	18.12	2.97***
Other	84.85	81.88	-2.97***
N	6,913	7,242	

The table reports for each categorical variable displayed in Table 4 the empirical frequencies of the A- and S-investors. The last column reports the difference between the A- and S-investors as well as the significance. *** indicates significance at 1%; ** indicates significance at 5%; * indicates significance at 10%.

matched sets with similar distributions of covariates, without requiring close or exact matches on all of the individual variables (Stuart (2010)).

To compute the propensity score for each investor in our sample, we build a logit model where the dependent variable is a binary variable that equals 1 if the investor has asked for an access to the information tool, and 0 otherwise.

As for the independent variables, Stuart (2010) suggest that matching using propensity scores must include variables associated both with the treatment (i.e. filling in the S-test) and the outcome (i.e. trading behavior). Our independent variables choice rest accordingly on two goals. First, we aim at including variables that are known to be associated with the demand for financial advice, i.e. financial literacy (Hackethal et al. (2012), Collins (2012), Georgarakos and Inderst (2014) and Calcagno and Monticone (2015)), education (Chalmers and Reuter (2010)) and socio-demographics (Bluethgen et al. (2008), Haslem (2008) and Gerhardt and Hackethal (2009)). In this respect, the level of financial literacy is controlled via the first variables of Table 4 on self-assessed knowledge and experience. The level of education and socio-demographic variables (gender, professional occupation, age) we control for are also found in the answers to the A-test.[18]

Second, we must also add variables that are associated with trading behavior. Again, socio-demographic variables have been documented to impact trading decisions and financial activity: women hold less risky assets (Bernasek and Shwiff (2001), Dwyer et al. (2002), Agnew et al. (2003), Charness and Gneezy (2012)) than men, age has an impact on the mix of risky assets, but it is not clear whether young invest more (Ackert et al. (2002)) or less in stocks (Shum and Faig (2006)). In addition, Grinblatt and Keloharju (2001) show that native tongue and culture have an impact on stockholdings and trades. In our data, as language indicates the main language of the investor (because of multilingual investors), we include it in the matching score determination. Financially literate households are known to invest more in stocks (Van Rooij et al. (2011), Balloch et al. (2014)) and in derivatives (Hsiao and Tsai (2018)). In fact, the variables that impact the demand for advice are also associated with trading behavior. Following Guiso and Jappelli (2006) we know that high trading frequency is observed for investors looking for information. Besides, according to Hackethal et al. (2010), investors who rely on advisors, have high trading volumes. As trading frequency and trading volumes are not always directly observed, trading activity is often proxied by portfolio size. Glaser (2003),

[18]For categorical variables, we only include N-1 dummies in the logit model. We consider the lowest level as the category of reference.

Vissing-Jorgensen (2003) and Abreu and Mendes (2012) have shown that portfolio size has indeed an impact on trading activity. Broihanne et al. (2016) also demonstrate that portfolio value is an important determinant of diversification choices. For example, Dorn and Huberman (2005), Guiso and Jappelli (2008), Goetzmann and Kumar (2008) find that investors with high portfolio size hold better diversified portfolios and earn higher returns. Finally, Nicolosi et al. (2009) show that the trading performance of individuals appears to improve with trading experience. For these reasons, our logit model also controls for portfolio value and trading experience.

Other papers on financial information and trading behavior build on matching procedure to create homogenous samples. Gerhardt and Hackethal (2009) use gender, age, marital status, risk tolerance, customer experience and deposit value to build comparative subsamples. Kramer (2012) matches on gender, age, residential value, income, portfolio and equity allocation. In contrast to our study, those papers investigate the effect of advice received from a financial advisor on trading behavior. To the best of our knowledge, the only paper that analyzes the effect of financial information acquisition on the trading behavior of retail investors using a similar database is Guiso and Jappelli (2006). While the latter identifies socio-demogaphic variables (wealth, risk tolerance and the level of education) as determinants of the search for financial information, they do not apply a random matching.[19] In addition, they do not investigate the effect of financial literacy.

Based on the propensity score, we associate for each S-investor the “closest” A-investor (“twin” A-investor) within a specific caliper width. According to Austin (2011), the optimal caliper width is 0.2 of the standard deviation of the logit of the propensity score. Since we have more S-investors than A-investors, we perform a random matching with replacement (i.e. one A-investor can be matched with several S-investors).

4 Results

4.1 Matching results

Table 6 reports the results of the logit model used to compute the propensity score. As for financial literacy, the investors who perceive themselves as highly literate and as expert in “complex” instruments are less likely to display an appetite for financial information. It is

[19] They instead use an instrumental variable approach.

to some extent consistent with Hung and Yoong (2010), Georgarakos and Inderst (2014) and Calcagno and Monticone (2015) who provide evidence that the self-perceived financial literacy is negatively correlated to the demand for financial advice. By contrast, the investors who stated to have effectively invested in "complex" instruments tend to be more information-seeker. While the opposite effect may seem surprising, this result is consistent with Calcagno and Monticone (2015).[20]

As for the level of education, the results suggest that the most educated investors tend to display a higher appetite for information. It is consistent with Hung and Yoong (2010), Collins (2012), Hackethal et al. (2012), and Hoechle et al. (2017) who report that the likelihood to ask for financial advice increases with the level of education. Focusing on the search for financial information, this result is also in line with Guiso and Jappelli (2006). They show that the most educated investors tend to spend more time looking for financial information. According to these authors, the level of education is a proxy for reduced cost of information.

As for the professional status, the investors who claim executive responsibilities tend to display a higher appetite for financial information. It is in contrast with Bluethgen et al. (2008) and Hackethal et al. (2012) who find no effect on the demand for financial advice. Furthermore, being a male significantly increases the appetite for information, which confirms the result observed in Guiso and Jappelli (2006). However, unlike previous studies, age does not seem to have any significant effect. Finally, the investors having a higher trading experience are more likely to fulfil the S-test and ask for more information. This is to some extent in line with Gerhardt and Hackethal (2009) while they show that the investors' ex ante trading experience is positively related to the decision to refer to a financial advisor.

Based on the propensity score, we build a sample of homogeneous investors. We end up with the 7,242 S-investors and 3,819 "matched" A-investors (i.e. an A-investor selected in the matching procedure has been matched on average with two S-investors). According to Stuart (2010), the standardized difference of means of the propensity score and the ratio of the variances of the propensity score in the treated and control group are the two most common numerical balance diagnostics to assess the quality of a matching. Rubin (2001) states that the standardized difference of means should be less than 0.25 and the variances ratio should be between 0.5 and 2. In our matched sample, the standardized difference of means of the propensity score is 0.13 while the variances ratio is 0.92, which indicates a correct matching

[20] They provide evidence that while subjective financial literacy is negatively correlated to the demand for financial advice, objective financial literacy has the opposite effect.

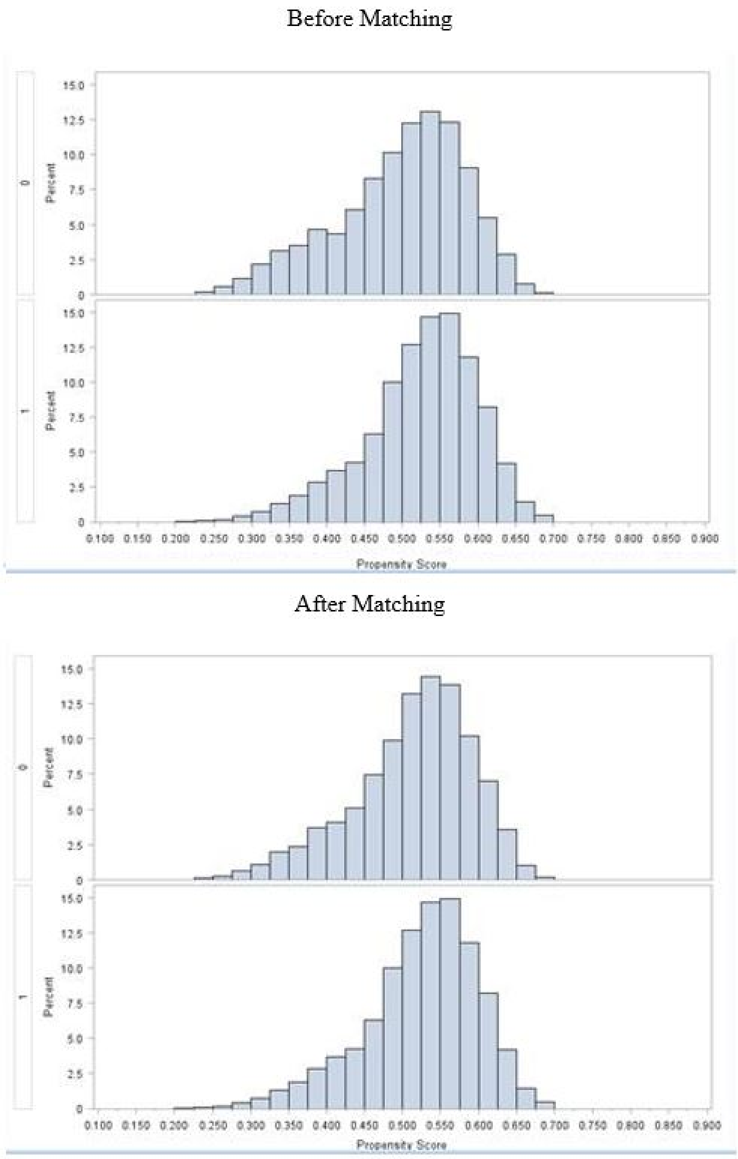
Table 6: Determinants of the appetite for information

Independent variables	Parameters estimates
Intercept	-1.0138***
Self-estimated knowledge of financial markets 1	-0.0671
Self-estimated knowledge of financial markets 2	-0.0532
Self-estimated knowledge of financial markets 3	-0.2697***
Self-evaluated experience in "complex" instruments 1	-0.2902***
Self-evaluated experience in "complex" instruments 2	-0.3251***
Investment in "complex" instruments "Yes"	0.1484***
Level of education 1	0.2121***
Level of education 2	0.3757***
Male	0.6137***
French-speaker	-0.1860***
English-speaker	-0.1798**
Executive	0.1366***
Age	-0.00106
Ln(PF value)	0.0174
Trading experience	0.00965***
Pseudo R^2	1.94%
N	14,155

The table reports parameters estimates of a logit model wherein the dependent variable is a binary variable that takes the value 1 if the investor has asked for an access to the information tool on stocks and has accordingly fulfilled the S-test; and 0 otherwise. The set of independent variables includes the variable presented in Table 4. It includes three dummies for the three highest levels of self-estimated knowledge of financial markets, two dummies for the two highest levels of self-evaluated experience in "complex" instruments, a dummy that takes the value 1 if the investor claims to have already invested in "complex" instruments, a dummy that takes the value 1 if the investor states to have a secondary/high school degree, a dummy that takes the value 1 if the investor states to have a university degree, a dummy that takes the value 1 if the investor is a male, a dummy that takes the value 1 if the investor is a French-speaker as his main language, a dummy that takes the value 1 if the investor is an English-speaker as his main language and a dummy that takes the value 1 if the investor claims executive responsibilities. In addition, the model also includes the age, the natural logarithm of the monthly average end-of-month portfolio market value and the trading experience. *** indicates significance at 1%; ** indicates significance at 5%; * indicates significance at 10%.

procedure. Figure 2 provides other evidence of the quality of our matching procedure. Distributions of the propensity score for the A- ("0") and S-investors ("1") before/after matching are reported. Figure 2 shows that the distribution of the propensity score of the A-investors after matching is closer from the one of the S-investors.

Figure 2: Propensity score distributions



The figure exhibits the distributions of the propensity score for the A- ("0") and S-investors ("1") before/after matching.

4.2 Trading behavior comparison results

In this section, we investigate whether the A- and S-investors display a different trading behavior through a univariate analysis. In this perspective, we distinguish three categories of variables that characterize trading behavior : trading activity, diversification and performance.

As measures of trading activity, we use the number of stock trades, the number of daytrades and the average number of trades on the same stock. We also consider whether investors trade options or warrants and bonds. Moreover, we consider the retail investors' exposure to the Disposition Effect (DE hereafter). This well-known behavioral bias refers to investors' reluctance to realize losses (i.e. they keep "losers") as well as their propensity to realize gains (i.e. they sell "winners"). In order to assess the DE at the individual level, we apply the methodology of Odean (1998a), by computing the difference between the proportion of gains realized and the proportion of losses realized.[21]

As measures of diversification, we build on Bellofatto et al. (2018) and use the number of different stocks traded during the whole period, that approximates an investor's investment universe, the monthly average number of stocks held in portfolio and the volatility of monthly returns. We also consider whether investors trade investment funds. As in Goetzmann and Kumar (2008) and Koestner et al. (2017), we consider the monthly average Herfindahl-Hirschman Index (HHI hereafter) too. Computed as the sum of squared stock portfolio weights, HHI is a measure of portfolio concentration that ranges from 0 (for well-diversified portfolios) to 1 (for underdiversified portfolios including only one stock). Building on Dorn and Huberman (2005) and Koestner et al. (2017), for investors who hold monthly positions in investment funds, we finally compute a modified Herfindahl-Hirschman Index (M_HHI hereafter) in which any position in funds is replaced by a portfolio of 50 equally-weighted securities.

As performance measures, we use the geometric average of both gross and net returns. We also consider gross and net Sharpe ratios as risk-adjusted measures of performance.

Table 7 reports the comparison of the trading behavior between the "matched" A- and S-investors. It exhibits for each variable the mean of both groups of investors, the difference between groups as well as the significance of the difference. As suggested by Stuart (2010), we

[21] For each investor, for each position in his portfolio, we compute on a daily basis his number of paper/realized gains and losses over his whole trading period. We then sum all those paper/realized gains and losses to compute his proportion of gains and losses realized.

take into account the number of times an A-investor has been selected as a match by using a frequency weights in the mean comparisons.

The results of Table 7 suggest that the A- and S-investors significantly differ in their trading behavior even after controlling for a large set of investors' characteristics. The differences in the volatility and the net Sharpe ratio are the only exceptions.

As for trading activity, while the S-investors make more trades, the A-investors make more daytrades and roundtrips on the same stock. Furthermore, the S-investors are more likely to invest in more "complex" instruments and display a lower exposure to the DE, which may suggest a higher level of financial experience and familiarity with financial markets (Bellofatto et al. (2018)).

As for diversification, the S-investors trade on a larger stock universe and tend to hold better diversified portfolios. Those investors include a higher number of stocks in their portfolios and are more likely to invest in investment funds. In addition, the HHI of the S-investors' portfolios is significantly lower. When taking into account the monthly holding in funds in the modified HHI, this finding is even more convincing. This finding is not in line with Guiso and Jappelli (2006) who report that the most information-seeker investors tend to have less diversified portfolios and tend to invest more in individual stocks. However, these authors use the proportion of the portfolio invested in funds as a proxy for diversification.

Finally, the S-investors display on average higher monthly gross and net returns than the A-investors. It is still at the opposite of Guiso and Jappelli (2008) who find a negative relationship between information acquisition and returns. On a risk-adjusted basis, this result only holds for the Gross Sharpe ratio.

Table 7: Univariate comparison results on trading behavior between "matched" A- and S-investors

	"matched" A-investors	S-investors	Difference
Number of stock trades	43.658	48.457	4.799***
Number of daytrades	1.6331	1.3694	-0.2637*
Average number of trades on the same stock	3.7095	3.1540	-0.5555***
Proportion of option traders (%)	17.16	19.41	2.25***
Proportion of bond traders (%)	2.39	4.03	1.64***
Disposition Effect (%)	17.7230	16.5495	-1.1735*
Number of different stocks traded	10.7211	13.6081	2.8870***
Number of stocks	3.7636	4.8258	1.0623***
Volatility (%)	18.5101	17.7121	0.7980
Proportion of fund traders (%)	16.09	27.51	11.42***
HHI (%)	58.05	51.00	-7.05***
Modified HHI (%)	58.05	45.99	-12.06***
Gross return (%)	0.172	0.506	0.334***
Net return (%)	-0.604	-0.338	0.266***
Gross Sharpe ratio	-0.00991	0.00684	0.0167***
Net Sharpe ratio	-0.0735	-0.0767	-0.00325
N	3,819	7,242	

The table reports for each variable the mean for the "matched" A- and S-investors, the difference between groups of investors as well as the significance of the difference. 'Number of stock trades' is the number of trades executed on stocks over the sample period. 'Number of daytrades' is the number of times an investor makes a purchase and a sale on the same stock on the same day over the sample period. 'Average number of trades on the same stock' is the average number of trades an investor makes on the same stock over the sample period. 'Proportion of option traders' is the proportion of investors who trade at least once either options or warrants over the sample period. 'Proportion of bond traders' is the proportion of investors who trade at least once bonds over the sample period. 'DE' refers to the disposition effect computed at the individual level as the difference between the proportion of gains realized and the proportion of losses realized. 'Number of different stocks traded' is the number of different stocks traded during the whole trading period. 'Number of stocks' is the monthly average number of stocks held in portfolio. 'Volatility' is the standard deviation of the monthly returns. 'Proportion of fund traders' is the proportion of investors who trade at least once investment fund shares over the sample period. 'HHI' is the monthly average Herfindahl-Hirschman Index, which is computed as the sum of squared stock portfolio weights. 'Modified HHI' is a modified version of the HHI for which any position in funds is replaced by a portfolio of 50 equally-weighted securities. 'Gross return' is the geometric average of the monthly gross returns. 'Net return' is the geometric average of the monthly net returns. 'Gross Sharpe-ratio' is a risk-adjusted measure of gross return. 'Net Sharpe-ratio' is a risk-adjusted measure of net return. *** indicates significance at 1%; ** indicates significance at 5%; * indicates significance at 10%.

5 Conclusion

In this paper, we use a unique opportunity offered in our data to test whether the investors who voluntarily ask for an information tool differ in their behavior from those who do not. Specifically, we investigate whether the investors displaying a distinct feature, that we call "appetite for information", trade differently than the investors neglecting an access to additional financial information. We identified this distinct feature through the mandatory MiFID tests. We find that the investors who display "appetite for information" invest smarter, i.e. trade funds, do less daytrades, are less subject to the DE, are better diversified and *in fine* earn higher returns, other trading and personal characteristics being controlled.

As for academia, our paper contributes to the literature on the relationship between trading behavior and information acquisition. While the vast majority of papers in this strand investigate this topic of research only descriptively, we base our analysis on actual trading records of about 14,000 retail investors over the 2008-2012 period. Furthermore, unlike precedent papers, we do not focus on trading frequency only but analyze trading behavior in a broader sense, including trading performance.

Our findings shed nevertheless light on two new perspectives that we need to discuss.

First, with those data at hand, we are not able to disentangle the effect of the distinct feature under study, the "appetite for information", from the effect of the information tool itself. The S-investors may trade differently than the A-investors because of their access to the information tool which asks for causality issues. While we cannot fully rule out this argument, it seems inconsistent with empirical evidence showing how access to more financial information impacts investors' behavior. In an important paper, Barber and Odean (2001b) report how the access to ongoing information has changed the trading behavior of retail investors. These authors suggest that abundant financial data, now easily accessible through the internet, has encouraged investors to trade more actively by bolstering overconfidence, fueled by an illusion of knowledge and control. In the same vein, Barber and Odean (2002) report consistent empirical evidence showing that after going online, investors tend to trade more actively and more speculatively. Recently, Benamar (2016) brings evidence that the introduction by a large brokerage house of a trade order management software (Trade+)[22] has significantly impacted the behavior of individual investors. This author finds that investors have started to implement

[22]The aim of the software is to display market data in a more efficient way that simultaneously gather all relevant information items (market data, centralized limit order book and investors' orders) into a user-customized screen

more aggressive trading strategies after the introduction of the software. Actually, the number of short-term trading strategies (short-term round-trips and daytrades) has increased for the investors who opted for this tool.

The previous papers have shown that, while a priori useful, additional information could lead to an illusion of knowledge that make investors believe that they have more abilities to trade. As a consequence, if the difference of trading behavior between the A- and S-investors was due to the information tool itself, we would expect the investors who get an access to the information tool to display more active and aggressive trading strategies as well as behavioral characteristics in line with the overconfidence theory. By contrast, the investors who ask for more financial information in our sample make less daytrades and less roundtrips. They also have a larger stock universe and include a higher number of stocks in their portfolios. In addition, it seems unlikely that the information tool under scrutiny explains by itself the difference in behavior in complex instruments since it represents a tool on stocks only.

A second line of discussion comes from the fact that the A-investors may decide to neglect the information tool on stocks, not because of a lack of appetite for information but because they already have an access to other sources of financial information. However, if the A-investors displayed appetite for information for other sources of information, it would be unlikely that they decide not to use the information tool provided by the investment firm under scrutiny. Indeed, this tool constitutes an additional (and valuable) means to get informed about the stocks.

In addition, our measure of "appetite for information" may share some characteristics with a personality trait[23] called "openness". Costa Jr and McCrae (1992) define it as the tendency of people to be open-minded and curious. According to personality psychologists, personality traits are key determinants of human behavior (Fung and Durand (2014)). Personality traits have recently been recognized as key variables explaining cross-sectional variations in investors' behavior. Several papers bring evidence that some personality traits are related to trading activity (Lo et al. (2005), Durand et al. (2008), Durand et al. (2013) and Tauni et al. (2015)). According to Costa Jr and McCrae (1992), individuals with high openness have favorable attitudes towards information and welcome it more easily in any context, whether this

[23]According to Roberts and Mroczek (2008), personality traits are defined as "*the relatively enduring patterns of thoughts, feelings, and behaviors that distinguish individuals from one another*". McCrae and Costa Jr (1997) support that a relatively small set of factors represents the basic dimensions of personality. There is a consensus among psychologists that the "Big Five" personality constructs of Norman (1963) best represents traits structure: (1) Neuroticism (N); (2) Extraversion (E); (3) Openness (O); (4) Agreeableness (A); and (5) Conscientiousness (C).

information has been searched out purposefully or encountered incidentally. These individuals also use imaginative and creative methods to acquire bulk information from a wide variety of sources of information (Kasperson (1978) and Palmer (1991)). As a consequence, since the investors displaying appetite for information are likely to be open-minded, we do not see any obvious reason that investors that have an access to other sources of information could choose to neglect an additional access to information about stocks.

Finally, our findings have implications for regulators (FED, ECB, ESMA) and investment firms. Our results suggest effectively that investors' behavior is consistent with their attitude towards financial information revealed through the MiFID tests. In line with their choice to neglect a free access to an information tool, the A-investors trade more intuitively while the investors displaying appetite for information tend to invest smarter. These findings are also important for assessing the usefulness of the MiFID questionnaire. By allowing to detect investors with appetite for information, MiFID questionnaires may contribute to the debate on the financial behavior of individual investors. While there is a consensus in the literature that there is a large heterogeneity across retail investors' financial behavior and performance (Barber and Odean (2013)), some factors have been recognized to play an important role: cognitive capacity (Christelis et al. (2010), Grinblatt et al. (2011)), trust (Guiso et al. (2008)), "sensitivity to the financial thing" (Guiso and Jappelli (2005)), the time spent collecting information (Guiso and Jappelli (2006)), social interactions (Hong et al. (2004)), optimism (Ben Mansour et al. (2006)), financial education (Van Rooij et al. (2011), Lusardi and Mitchell (2014)) or even "happiness" (Kaplanski et al. (2015)). We hope that our findings add a new qualitative determinant, namely the "appetite for information".

References

Abreu, M. and V. Mendes (2012). Information, overconfidence and trading: Do the sources of information matter? *Journal of Economic Psychology 33*(4), 868–881.

Ackert, L. F., B. K. Church, and B. Englis (2002). The asset allocation decision and investor heterogeneity: a puzzle? *Journal of economic behavior & organization 47*(4), 423–433.

Agnew, J., P. Balduzzi, and A. Sunden (2003). Portfolio choice and trading in a large 401 (k) plan. *American Economic Review 93*(1), 193–215.

Argentesi, E., H. Lütkepohl, and M. Motta (2010). Acquisition of information and share prices: An empirical investigation of cognitive dissonance. *German Economic Review 11*(3), 381–396.

Austin, P. C. (2011). Optimal caliper widths for propensity-score matching when estimating differences in means and differences in proportions in observational studies. *Pharmaceutical statistics 10*(2), 150–161.

Balloch, A., A. Nicolae, and D. Philip (2014). Stock market literacy, trust, and participation. *Review of Finance 19*(5), 1925–1963.

Barber, B. M. and T. Odean (2000). Trading is hazardous to your wealth: The common stock investment performance of individual investors. *Journal of Finance 55*(2), 773–806.

Barber, B. M. and T. Odean (2001a). Boys will be boys: Gender, overconfidence, and common stock investment. *Quarterly Journal of Economics 116*(1), 261–292.

Barber, B. M. and T. Odean (2001b). The internet and the investor. *Journal of Economic Perspectives 15*(1), 41–54.

Barber, B. M. and T. Odean (2002). Online investors: Do the slow die first? *Review of Financial Studies 15*(2), 455–488.

Barber, B. M. and T. Odean (2013). The behavior of individual investors. In *Handbook of the Economics of Finance, volume 2, part B, edited by George M. Constantinides, Milton Harris and Rene M. Stulz, 1533-1570.*

Bellofatto, A., C. D'Hondt, and R. De Winne (2018). Subjective financial literacy and retail investors' behavior. *Journal of Banking & Finance 92*, 168–181.

Ben Mansour, S., E. Jouini, and C. Napp (2006). Is there a “Pessimistic” bias in individual beliefs? Evidence from a simple survey. *Theory and Decision 61*(4), 345–362.

Benamar, H. (2016). To see is to know: Simultaneous display of market data for retail investors. *Available at SSRN: https://ssrn.com/abstract=2336035*.

Bernasek, A. and S. Shwiff (2001). Gender, risk, and retirement. *Journal of economic issues 35*(2), 345–356.

Bluethgen, R., A. Gintschel, A. Hackethal, and A. Mueller (2008). Financial advice and individual investors' portfolios. *Available at SSRN 968197*.

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

Broihanne, M.-H. and H. Orküt (2018). Investment goals and mental accounting in french retail clients. *Finance (Forthcoming)*.

Bucher-Koenen, T. and J. Koenen (2010). Do smarter consumers get better advice? An analytical framework and evidence from german private pensions. *CDSE Discussion Paper No. 105*.

Calcagno, R. and C. Monticone (2015). Financial literacy and the demand for financial advice. *Journal of Banking & Finance 50*, 363–380.

Chalmers, J. and J. Reuter (2010). What is the impact of financial advisors on retirement portfolio choices and outcomes? *National Bureau of Economic Research (NB10-05)*.

Charness, G. and U. Gneezy (2012). Strong evidence for gender differences in risk taking. *Journal of Economic Behavior & Organization 83*(1), 50–58.

Christelis, D., T. Jappelli, and M. Padula (2010). Cognitive abilities and portfolio choice. *European Economic Review 54*(1), 18–38.

Collins, J. M. (2012). Financial advice: A substitute for financial literacy? *Financial Services Review 21*(4), 307–322.

Committee of European Securities Regulators (2008). A consumer's guide to MiFID: Investing in financial products.

Costa Jr, P. T. and R. R. McCrae (1992). Normal personality assessment in clinical practice: The neo personality inventory. *Psychological Assessment 4*(1), 5–13.

D'Hondt, C. and P. Roger (2017). Investor sentiment and stock return predictability: The power of ignorance. *Finance 38*(2), 7–37.

Dorn, D. and G. Huberman (2005). Talk and action: What individual investors say and what they do. *Review of Finance 9*(4), 437–481.

Durand, R., R. Newby, K. Tant, and S. Trepongkaruna (2013). Overconfidence, overreaction and personality. *Review of Behavioral Finance 5*(2), 104–133.

Durand, R. B., R. Newby, and J. Sanghani (2008). An intimate portrait of the individual investor. *Journal of Behavioral Finance 9*(4), 193–208.

Dwyer, P. D., J. H. Gilkeson, and J. A. List (2002). Gender differences in revealed risk taking: evidence from mutual fund investors. *Economics Letters 76*(2), 151–158.

Fung, L. and R. B. Durand (2014). Personality traits. In *Investor behavior: The psychology of financial planning and investing, edited by H.K. Baker, & V. Ricciardi (Eds.), 99-155, Hoboken, NJ: John Wiley & Sons, Inc.*

Georgarakos, D. and R. Inderst (2014). Financial advice and stock market participation. *Working Paper No 1296, European Central Bank*.

Gerhardt, R. and A. Hackethal (2009). The influence of financial advisors on household portfolios: A study on private investors switching to financial advice. *Available at SSRN 1343607*.

Glaser, M. (2003). Online broker investors: Demographic information, investment strategy, portfolio positions, and trading activity. *Available at SSRN 975985*.

Glaser, M. and M. Weber (2007). Overconfidence and trading volume. *Geneva Risk and Insurance Review 32*(1), 1–36.

Glaser, M. and M. Weber (2009). Which past returns affect trading volume? *Journal of Financial Markets 12*(1), 1–31.

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

Grinblatt, M. and M. Keloharju (2001). How distance, language, and culture influence stockholdings and trades. *Journal of Finance 56*(3), 1053–1073.

Grinblatt, M., M. Keloharju, and J. Linnainmaa (2011). IQ and stock market participation. *Journal of Finance 66*(6), 2121–2164.

Grossman, S. J. and J. E. Stiglitz (1980). On the impossibility of informationally efficient markets. *American Economic Review 70*(3), 393–408.

Guiso, L. and T. Jappelli (2005). Awareness and stock market participation. *Review of Finance 9*(4), 537–567.

Guiso, L. and T. Jappelli (2006). Information acquisition and portfolio performance. *CSEF Working Paper No. 167.r*.

Guiso, L. and T. Jappelli (2008). Financial literacy and portfolio diversification. *Working Paper, European University Institute*.

Guiso, L., P. Sapienza, and L. Zingales (2008). Trusting the stock market. *Journal of Finance 63*(6), 2557–2600.

Hackethal, A., M. Haliassos, and T. Jappelli (2012). Financial advisors: A case of babysitters? *Journal of Banking and Finance 36*(2), 509–524.

Hackethal, A., R. Inderst, and S. Meyer (2010). Trading on advice. *CEPR Discussion Paper No. DP8091*.

Haslem, J. A. (2008). Why do mutual fund investors employ financial advisors? *Journal of Investing 17*(4), 91–94.

Hoechle, D., S. Ruenzi, N. Schaub, and M. Schmid (2017). The impact of financial advice on trade performance and behavioral biases. *Review of Finance 21*(2), 871–910.

Holthausen, R. W. and R. E. Verrecchia (1990). The effect of informedness and consensus on price and volume behavior. *Accounting Review 65*(1), 191–208.

Hong, H., J. D. Kubik, and J. C. Stein (2004). Social interaction and stock-market participation. *Journal of finance 59*(1), 137–163.

Hsiao, Y.-J. and W.-C. Tsai (2018). Financial literacy and participation in the derivatives markets. *Journal of Banking & Finance 88*, 15–29.

Hung, A. and J. Yoong (2010). Asking for help: Survey and experimental evidence on financial advice and behavior change. *RAND Working Paper Series WR-714-1*.

Kaplanski, G., H. Levy, C. Veld, and Y. Veld-Merkoulova (2015). Do happy people make optimistic investors? *Journal of Financial and Quantitative Analysis 50*(1-2), 145–168.

Karpoff, J. M. (1986). A theory of trading volume. *Journal of Finance 41*(5), 1069–1087.

Kasperson, C. J. (1978). Psychology of the scientist: Xxxvii. Scientific creativity: A relationship with information channels. *Psychological Reports 42*(3), 691–694.

Koestner, M., B. Loos, S. Meyer, and A. Hackethal (2017). Do individual investors learn from their mistakes? *Journal of Business Economics 87*(5), 669–703.

Kramer, M. M. (2012). Financial advice and individual investor portfolio performance. *Financial Management 41*(2), 395–428.

Kumar, A. and C. Lee (2006). Retail investors sentiment and returns comovements. *Journal of Finance 61*(5), 2451–2486.

Lo, A. W., D. V. Repin, and B. N. Steenbarger (2005). Fear and greed in financial markets: A clinical study of day-traders. *American Economic Review 95*(2), 352–359.

Lusardi, A. and O. S. Mitchell (2014). The economic importance of financial literacy: Theory and evidence. *Journal of Economic Literature 52*(1), 5–44.

Mazzoli, C. and N. Marinelli (2014). Determinants of risk-suitable investment portfolios: Evidence from a sample of italian householders. *International Journal of Finance & Banking 1*(1), 13–28.

McCrae, R. R. and P. T. Costa Jr (1997). Personality trait structure as a human universal. *American Psychologist 52*(5), 509.

Nicolosi, G., L. Peng, and N. Zhu (2009). Do individual investors learn from their trading experience? *Journal of Financial Markets 12*(2), 317–336.

Norman, W. T. (1963). Personality measurement, faking, and detection: An assessment method for use in personnel selection. *Journal of Applied Psychology 47*(4), 225.

Odean, T. (1998a). Are investors reluctant to realize their losses? *Journal of finance 53*(5), 1775–1798.

Odean, T. (1998b). Volume, volatility, price, and profit when all traders are above average. *Journal of Finance 53*(6), 1887–1934.

Palmer, J. (1991). Scientists and information: II. Personal factors in information behaviour. *Journal of documentation 47*(3), 254–275.

Polkovnichenko, V. (2010). Individual investor portfolios. In *Behavioral Finance: Investors, Corporations, and Markets (Robert W. Kolb Series in Finance), edited by H. Kent Baker and John R. Nofsinger, 539–557. John Wiley & Sons.*

Roberts, B. W. and D. Mroczek (2008). Personality trait change in adulthood. *Current Directions in Psychological Science 17*(1), 31–35.

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

Rubin, D. B. (2001). Using propensity scores to help design observational studies: application to the tobacco litigation. *Health Services and Outcomes Research Methodology 2*(3-4), 169–188.

Shu, P.-G., S.-B. Chiu, H.-C. Chen, and Y.-H. Yeh (2004). Does trading improve individual investor performance? *Review of Quantitative Finance and Accounting 22*(3), 199–217.

Shum, P. and M. Faig (2006). What explains household stock holdings? *Journal of Banking & Finance 30*(9), 2579–2597.

Stuart, E. A. (2010). Matching methods for causal inference: A review and a look forward. *Statistical science 25*(1), 1–21.

Tauni, M. Z., H. X. Fang, S. Yousaf, et al. (2015). The influence of investor personality traits on information acquisition and trading behavior: Evidence from chinese futures exchange. *Personality and Individual Differences 87*, 248–255.

Van Rooij, M., A. Lusardi, and R. Alessie (2011). Financial literacy and stock market participation. *Journal of Financial Economics 101*(2), 449–472.

Vissing-Jorgensen, A. (2003). Perspectives on behavioral finance: Does "irrationality" disappear with wealth? Evidence from expectations and actions. In *NBER Macroeconomics Annual 2003, edited by M. Gertler and K. Rogoff, 139-208.* The MIT Press.