EIEF WORKING PAPER SERIES

Einaudi Institute for Economics and Finance

EIEF Working Paper 18/07
June 2018

Investment in Financial Information and Portfolio Performance

by

Luigi Guiso
(EIEF and CEPR)

Tullio Jappelli
(University of Naples Federico II, CSEF and CEPR)

Investment in Financial Information and Portfolio Performance

Luigi Guiso
EIEF and CEPR.

Tullio Jappelli
University of Naples Federico II, CSEF and CEPR.

June 2018

Abstract

Rational investors perceive correctly the value of financial information. Investment in information is therefore associated with higher expected portfolio returns and Sharpe ratio. Overconfident investors overstate the quality of their own information, and thus investment in information is associated with a lower expected Sharpe ratio despite they realize higher average returns. We contrast the implications of these two models using two unique surveys of customers of a leading Italian bank with portfolio data and measures of financial information. We find that the investment in information is positively associated with returns to financial wealth and negatively to Sharpe ratio. The latter falls with proxies for overconfidence. We relate these findings to the wealth inequality debate.

JEL: E2, D8, G1
Keywords: Portfolio Choice, Information, Overconfidence

We thank Giovanni Cespa, Andrea Frazzini, Uri Gneezy, Nick Souleles and Joel Peress for helpful discussions on a previous version of the paper, and seminar participants at the American Economic Association, University of Chicago Graduate School of Business, Dartmouth College, University of Turin, and University of Siena for comments. Raghu Suryanarayanan and Veronica De Falco provided excellent research assistance. We are grateful to Unicredit, and particularly to Daniele Fano and Laura Marzorati, for granting us access to the UCS surveys. This work has been partly supported by the Italian Ministry of Education, Universities and Research (MIUR).

1. Introduction

How much financial information should investors collect? And what is the effect of information on portfolio performance? These questions are currently at the heart of the debate about the determinants of wealth inequality that emphasizes the critical importance of heterogeneity in returns to wealth, see Benhabib and Bisin (2018) for a recent review. And heterogeneity in investors information has, since Arrow (1987), been argued to be a potentially key ingredient in generating heterogeneity in returns (Best and Dogra, 2017; Kacperczyk et al. 2018).

In models with rational investors the answer to the above questions is straightforward. Investors should spend time and money collecting financial information up to the point where the marginal benefit of doing so exceeds the marginal cost. Since investors acquire more information only if their utility increases, information improves portfolio performance. Indeed, Peress (2004), building on the seminal work by Arrow (1987), shows that the portfolio expected return and Sharpe ratio of rational agents increase with the amount of information they optimally collect.

But other models deliver different predictions. Drawing on a large body of evidence from experimental cognitive and psychological research, one class of models argues that many investors are overconfident when they make financial decisions. Overconfident investors collect too much private information, trade more and take more risk than agents with unbiased perceptions. As a result, they may earn higher average portfolio returns as compensation for risk but attain poorer portfolio performance, as measured by the Sharpe ratio. In a model with endogenous information acquisition, Odean (1998a) shows that overconfident investors are more likely to be informed and obtain lower utility than rational investors who choose to remain uninformed. Using a survey of accounts at a discount broker, Odean (1998b) and Barber and Odean (1999, 2001) find that investors make unprofitable trades in the sense that the assets they buy tend, on average, to under-perform the assets they sell, resulting in negative profits from trading even before trading costs are accounted for. In addition, men - arguably more overconfident than women according to the experimental psychology literature - trade more often and therefore perform less well than women.[1]

The hypothesis underlying the overconfidence model is that investors overestimate the value of the private signals and, for this reason, spend too much money and time acquiring information. In turn, overconfidence leads to inefficient portfolio allocations and trades, the more so the more information is acquired. This suggests that a proper test of this departure from rationality requires data on financial information and portfolio performance.

In this paper we provide such test. We contrast the rational and overconfidence models studying the determinants of information acquisition and the correlation between information and portfolio performance. Peress (2004) shows that the portfolio Sharpe ratio of rational investors - who maximize expected utility and process information correctly - is positively correlated with the amount of private information acquired. Indeed, it is precisely the expected benefit of attaining a higher Sharpe ratio that induces investors to incur the cost of acquiring information.

Overconfident investors face the same incentives. But given that they overestimate the value of information, the Sharpe ratio they obtain is lower than the Sharpe ratio they think they would obtain based on the wrong assessment of the precision of their information. Most importantly, we show that if investors are sufficiently overconfident, their portfolio Sharpe ratio is negatively correlated with the amount of information they collect, and that this negative correlation is stronger when overconfidence increases. Our test distinguishes the two models relying on variables that are observable and measurable.

To implement the test, we use data from two surveys of investors randomly sampled from customers of a leading Italian bank, with data on time people spend acquiring financial information, risk attitudes, trading and socioeconomic variables. Detailed financial data allow us to construct a measure of the portfolio expected return and volatility for each investor. Given the assumptions required to estimate the expected

[1] Biais et al. (2005) reach similar conclusions in an experimental setting, where they relate directly trading performance to a measure of overconfidence obtained independently as part of the experiment.

Sharpe ratio, we check the robustness of the results using alternative measures constructed from historical returns and volatility of the actual investors' portfolios.

In a first part of our analysis, we find that investment in financial information increases with wealth and risk tolerance, and is negatively associated with proxies of the cost of information. The findings are consistent with both the rational and the overconfidence models, as both predict that investors who benefit more from extra information (the wealthy and the risk tolerant, because they invest more in risky and information intensive assets) and those who can obtain information at lower cost, collect more information. This evidence suggests that investors respond strongly to economic incentives in deciding how much information they acquire.

In a second step, we find that the average portfolio return is positively associated with investment in information as predicted by both the rational and overconfidence model. But the constructed proxies for the expected Sharpe ratio are negatively associated with investment in information, consistent with the overconfidence model. The relation is unchanged if we add further controls, and is robust to different sample definitions and sample selection. The relation we find is also economically important: in our baseline estimates those who spend between 2 and 4 hours per week in acquiring financial information obtain a 22 basis points higher average portfolio return (20 percent of the sample average portfolio return) and have a Sharpe ratio that is 27 percent lower than those who spend no time. Evaluated at the sample median of the portfolio standard deviation, this is equivalent to a reduction of 16 basis points in the portfolio expected excess return.

The negative relation between the Sharpe ratio and information might be driven by unobserved heterogeneity, for instance because those who enjoy trading stocks - a utility benefit that does not show up in measured portfolio returns and is not observed by the econometrician - also enjoy collecting more financial information. We address this issue by an instrumental variables approach, using as instruments variables that are unlikely to be related with preference or taste for finance. In the instrumental variables regressions the negative relation between information and the Sharpe ratio is, if anything, reinforced.

Furthermore, the negative relation is stronger for groups that are, a priori, expected to be more overconfident.

Overall, our evidence conflicts with the fully rational model, and supports models where investors overstate the quality of information, invest too much in information and take excessive financial risk. While these conclusions are similar to Odean (1998b) and Barber and Odean (2001), there are important substantive and methodological differences. First, our surveys allow us to contrast the predictions of the rational and overconfidence models from quite different perspectives, increasing the robustness and reliability of the results. Second, our results rely on representative samples of retail investors with a bank account, while previous studies focus mostly on samples of investors at discount brokers. These are highly selected samples of investors who trade stocks directly.[2] Thus, they are likely to include relatively more investors with a predisposition for overconfidence or who are willing to incur losses for the pleasure of trading. Discriminating between these two alternatives is not easy with Odean's (1998b) data. Our instrumental variables approach allows us to rule out the second possibility, while our representative sample limits the sample selection problem. Finally, while Odean (1998b) and Barber and Odean (2001) administrative data focus on common stock trading, we look at the performance of the entire financial portfolio.

Our paper is closely related to a burgeoning literature that uncovers substantial heterogeneity in portfolio returns and Sharpe ratios. The evidence comes from detailed analysis of portfolio performance in Scandinavian countries for which extensive panel data on individual accounts are available. Calvet et al. (2007, 2009) find considerable heterogeneity in financial portfolio performance using Swedish data, and show that proxies of financial sophistication (such as wealth, income, occupation and education) are associated with higher Sharpe ratios. Bach et al. (2017) using an administrative panel of all Swedish residents, document that returns on financial wealth are on average 4% higher for households in the top 1% compared to the median household. Fagereng et al.

[2] Bilias et al. (2010) study households' portfolio inertia using data from the PSID and the Survey of Consumer Finances. In these representative samples households seldom trade: over a 5-year period (1994-99), 73.8 percent did not trade stocks. This contrasts with the trading activity of a minority of investors (less than 20 percent) that have a brokerage account (not necessarily a discount account): 70 percent trades stocks at least once a year.

(2018) use 15 years of population data from Norway's administrative tax records and find that individuals earn markedly different returns on their financial assets, with a difference of 300 basis points between the 10th and the 90th percentile, returns are heterogeneous even within narrowly defined asset classes and heterogeneity extends to average observed Sharpe ratios. Furthermore, as in Calvet et al. (2007, 2009), they also document that wealth and financial sophistication predict portfolio Sharpe ratios.

Like these recent papers, we also document significant heterogeneity in the Sharpe ratio of financial portfolios. But we enrich the interpretation by showing that heterogeneity in the portfolio Sharpe ratio can be traced back to differences in consumers financial information and in the incentives to invest in information, as emphasized by Arrow (1987) and studied more recently by Best and Dogra (2017) and Kacperczyk et al. (2018). We add to heterogeneity in information in explaining portfolio performance, also heterogeneity in investors' ability to appreciate the value of information, i.e. overconfidence.

A related line of research studies stock concentration. For instance, Huberman (2001) and Boyle et al. (2012) find that portfolios are more concentrated in stocks that people are more familiar with, and that there are "returns to concentration." Massa and Simonov (2006), using Swedish administrative data, find that concentrated stocks are those to which the investor is geographically or professionally closer, or that he or she has held for a long time. Ivkovic et al. (2008), using data from a US discount broker, find that portfolios with concentrated stocks actually outperform more diversified accounts. One potential explanation of these findings is that investors with concentrated portfolios are able to exploit some informational advantage that allows them to pick up winning stocks, as argued by Van Nieuwerburgh and Veldkamp (2010). This is only a conjecture, however, because in these studies investors' information is not observed. On this front, we find that investors who acquire more information tend to have less diversified portfolios; but at the same time those who diversify less attain a lower Sharpe ratio.

The rest of the paper proceeds as follows. Section 2 reviews the theoretical literature, contrasting the predictions of the rational and overconfident models for the relation between investment in information and portfolio performance, and summarizes the main insight of a theoretical model that is detailed in Appendix A. Section 3 describes the survey and explains how we measure investment in information and portfolio performance. Section 4 presents evidence on the determinants of information acquisition. Section 5 presents the main results of the paper, relating the Sharpe ratio to investment in information. Section 6 summarizes the results.

2. Theoretical framework

In models with rational investors the set of variables that affect asset allocation and information acquisition are well identified. In a seminal paper Verrecchia (1982) shows that investors with higher cost of acquiring information and risk averse investors acquire less information, the latter because they intend to invest less in stocks and therefore information is less valuable for them. These empirical predictions, however, don't discriminate between rational and overconfident investors. Indeed, as we will see, overconfident investors behave very much like rational investors with respect to the determinants of information acquisition. However, the implications for the effect of information on portfolio performance are different.

Peress (2004), building on seminal work by Arrow (1987), shows that in a model with rational investors information improves the allocation of wealth and is associated with a higher expected Sharpe ratio. Although the portfolio of informed investors is riskier and thus earns higher expected returns, the risk-adjusted return is higher. In contrast, overconfident investors acquire more information and react to information more strongly than rational investors. As in the rational model, portfolio risk and return increase with information. But the Sharpe ratio of the portfolio of an overconfident investor may be lower. This section presents a framework to distinguish the two models empirically. We summarize here the main theoretical propositions, and show the details in Appendix A.

2.1. Rational investors

Starting with Grossman and Stiglitz (1980), several papers propose models of rational investors where agents can increase, at a cost, information on the random return of a risky asset, see Verrecchia (1982), Barlevy and Veronesi (1999), Van Nieuwerburgh and Veldkamp (2010), Mondria (2010), and Kacperczyk et al (2018). Drawing on Peress (2004), Appendix A presents a model that delivers several testable predictions.

First, information purchased increases with investor's wealth and risk tolerance, and falls with the marginal cost of information. Wealthier and more risk tolerant investors value information more because they invest more wealth in the information intensive asset and, accordingly, the signal is more valuable for them. Second, corner solutions can be optimal. Poor or very risk averse investors benefit little from information, because they would invest little in stocks even if they had a very precise signal. Thus, they may choose to purchase no information. Third, the expected portfolio return and volatility increase with information. More informed investors face less risk and invest more aggressively in stocks, obtaining higher returns. They react also more strongly to the signals they receive and trade more.

The fourth implication of the model is that rational agents are willing to pay the cost of information precisely because they expect to obtain a benefit in terms of higher risk-adjusted return. This implies that the expected Sharpe ratio increases with information purchased, even accounting for trading and information costs. Finally, risk aversion affects the Sharpe ratio only because it affects information purchased. In other words, risk aversion should not affect the Sharpe ratio, holding information constant. This is a neat exclusion restriction of the rational model that we are able to confront with the data.

2.2. Overconfident investors

Overconfident investors maximize expected utility, like rational investors. But unlike rational investors, overconfident investors overestimate the signal's precision.[3] In

[3] Here is one of many examples of overconfidence. In 1991, the US General Social Survey asked the following two questions: (1) "Compared to other people who do the same or similar kind of work that you do, how well would you say you do your job? Would you say much better, somewhat better, about the same, somewhat worse or much worse?" (2) "Compared to other people who do the same or similar kind of

the model presented in Appendix A the decision to purchase information is driven by the same variables as in the rational model: wealth, risk tolerance and cost of information. However, overconfident investors purchase more information because the perceived value of information is higher than its true value. Proxies for overconfidence - for instance gender as in Lundeberg et al. (1994) and Barber and Odean (2001) - should therefore help predicting investment in information. But apart from this, the information decision of overconfident investors is observationally equivalent to that of rational agents. This implies that the determinants of investment in information alone do not allow discriminating between the rational and the overconfidence model.

The difference between the two models lies in the consequence of information on portfolio performance. Odean (1998a) shows that overconfident investors attain lower utility than rational investors and take more risk, for given expected return, and attain a lower Sharpe ratio. But we go beyond this result, showing that overconfidence affects the relationship between portfolio performance and the amount of information purchased. Our main results are summarized in the following two propositions:

(1) If investors are sufficiently overconfident, the expected Sharpe ratio, obtained conditioning on the true signal rather than on the perceived signal, is a decreasing function of investment in information. Proof: See Appendix A.

(2) The more investors are overconfident, the more negative is the relation between the expected Sharpe ratio and investment in information. Proof: See Appendix A.

The first proposition predicts a relation between information and the Sharpe ratio opposite to that implied by the rational model, at least for high levels of overconfidence. This suggests that one can discriminate between the two models using variables that are observable and measurable, at least potentially. Our second proposition predicts different

work that you do, how much work would you say you do? Would you say that you do much more, somewhat more, about the same, somewhat less or much less?" Over 72% percent answered to the first question they did better or much better than average; only 0.2% rated themselves below average. About 61% said they worked more or much more than other people, and only 3.3% below average.

slopes of the relation between the Sharpe ratio and information in groups of investors with low or high overconfidence. Indeed, empirical research shows that overconfidence depends on specific domains of activities as well as individual attributes.[4]

Table 1 summarizes the empirical predictions of the rational and overconfident models. In both models, the numerator and the denominator of the Sharpe ratio increase with the amount of information. However, in the overconfident model one extra unit of information raises the standard deviation of the portfolio more than its expected return, because the misperception of the signal's precision induces investors to take some uncompensated risk.

3. Data description

The Unicredit Survey of Investors Behavior (UCS) is a very detailed survey of 1,834 customers of Unicredit, a leading Italian commercial bank with over 4 million accounts. The sample is representative of the population of Unicredit retail customers with a bank account (whereas 15% of the Italian population has not) and refers to 2004. Unicredit has a large market share, and relatively more customers in Northern Italy where people are wealthier on average. The UCS therefore tends to over-sample relatively rich investors. The unit of observation is the customer, defined as a person with an account in one of Unicredit banks. Appendix B describes sample design and other characteristics of the survey.

Differently from other investors' surveys, UCS asks investors to report information on real and financial assets of all household members, both inside and outside Unicredit. It collects data on investment in financial information, knowledge of specific financial assets, attitudes towards financial risk, bank-customer relations, and reliance on financial advice. The UCS represents therefore a unique opportunity to study the relation between

[4] Overconfidence can be substantial especially when people face range questions. For instance, Russo and Schoemaker (1992) find that businessmen asked to provide 90 percent confidence ranges have the correct answer within the stated range only 42 to 62 percent of the time; Klayman et. al. (1999) find similar results in an experiment that accounts for confounding statistical effects when measuring overconfidence.

financial information, portfolio allocation and portfolio performance, and to confront with the data the implications of the rational and overconfidence models outlined in Section 2.

For robustness, we also use a second survey, taken in 2007. Differently from the 2004 survey it has no information on assets outside Unicredit, but it has detailed panel administrative data from 2004 to 2007 on financial wealth at Unicredit of all retail investors surveyed in 2004. These additional results are reported in Section 5.3.

3.1. Investment in financial information

The UCS has a question on time spent acquiring financial information: “Let's talk about financial information. How much time do you usually spend, in a week, to obtain information on how to invest your savings? (think about time reading newspapers, surfing the internet, talking to your advisor, reading companies balance sheets, etc.).”

Answers range from no time to more than 7 hours per week. Table 2 displays the sample distribution of the variable. Over one third of the sample spends no time, most respondents spend “Less than 30 minutes” or “Between 30 and 60 minutes” per week. At the other extreme, 13 percent of the sample spends more than 2 hours per week (5 percent of the average weekly working time). To provide further insights on the amount of time involved, the last row of Table 2 reports the equivalent number of working days spent in information each year. The number ranges from zero to 43 days, documenting substantial heterogeneity in the time investors spend gathering financial information.

As suggested in Section 2, in both the rational and overconfident models those who invest in stocks have a stronger incentive to acquire information than those who don’t. On the other hand, those who are more informed perceive lower return volatility and have less incentives to invest more in stocks. Consistent with these predictions those who collect more information are also more likely to own stocks and to invest a larger share of their wealth in stocks.[5]

3.2. Financial wealth and portfolio performance

Financial wealth is constructed from questions on ten different assets categories: (1) bank accounts; (2) repurchase agreements; (3) certificate of deposits; (4) government bonds; (5) corporate bonds; (6) derivatives: (7) shares of listed companies; (8) shares of unlisted companies; (9) mutual funds; (10) managed investment accounts. For each of these categories, the survey provides information on assets kept with Unicredit, as well as with other banks and financial institutions. Total financial wealth is the sum of all financial assets, both in Unicredit accounts and in other banks and financial institutions. Two definitions of financial wealth are available: respondents' wealth (the bank's customer), and household financial wealth, resulting from the sum of respondent' and other household members wealth, see Appendix B for details.

Our measure of expected returns and volatility is based on the same procedure and assumptions as in Pelizzon and Weber (2005). We combine survey information on the ten financial assets with time series data on assets returns and compute, for each investor, the portfolio expected return and volatility, as described in Appendix B.

Since not all investors own risky assets, the Sharpe ratio is defined for 1,365 out of 1,834 observations, 74.4% of the total sample. The remaining part of the sample invests only in risk-free assets. The average Sharpe ratio is estimated at 0.26. In contrast to the uniformity of the Sharpe ratio predicted by standard finance theory, the observed ratio exhibits considerable sample variability, ranging from 0.108 to 0.538 with a standard deviation is 0.15.

[5] Stock market participation is positively correlated with investment in information but the direction of causality is not obvious. If investors choose information after the participation decision, those who don't participate should not purchase information (unless they do it for pleasure). If information is purchased before the participation decision, some who don't participate may have purchased information, but have chosen to stay out of the market on the basis of the information purchased. In the data, even among those who acquire information, some don't buy stocks, suggesting that information is acquired before the participation decision, at least for this group.

3.3. Risk aversion

The UCS has an indicator of risk aversion patterned after the Survey of Consumer Finance: "Which of the following statements comes closest to the amount of financial risk that you are willing to take when you make your financial investment: (1) a very high return, with a very high risk of losing the money; (2) high return and high risk; (3) moderate return and moderate risk; (4) low return and no risk."[6]

Only 19% choose "low return and no risk", so most are willing to accept some risk if compensated by a higher return. A recent literature on eliciting preferences from survey data shows that direct questions on risk aversion are informative and have predictive power.[7]

The survey has also another indicator of risk aversion obtained from the question: "With which of the following statements do you agree most? (1) Risk is an uncertain event from which one can extract a profit; (2) Risk is an uncertain event from which one should seek protection." Most respondents (71%) answer (2), considering risk a threat rather than an opportunity. The two indicators of risk aversion, though based on quite different framing, are highly correlated. In the empirical analysis we rely mostly on the first indicator, but check the sensitivity of the results using also the second. Table 3 reports sample statistics for the risk aversion indicators.

Finally, the UCS also has detailed socioeconomic variables for the respondent and household members: education, gender, marital status, and residence. Summary statistics for the variables used in the estimation are also reported in Table 3.

4. Determinants of investment in financial information

As shown in Section 2, the rational model and the model with overconfident but utility-maximizing investors deliver similar predictions on the determinants of investment

[6] The question does not distinguish between relative and absolute risk aversion. But since we can control for wealth, we can allow the risk aversion indicator to reflect differences in risk preferences that don't arise from differences in endowments.

[7] See, among others, Guiso and Paiella (2008) and Dohmen et al (2011).

in information. Thus, one cannot rely on estimates of the demand for financial information alone to discriminate between the two models. Yet, looking at these determinants is useful for several reasons.

First, uncovering the determinants of financial information acquisition matters considerably for models of wealth inequality that hinge on heterogeneous incentives to acquire information as a source of heterogeneity in returns to wealth, and thus a cause of wealth concentration. Second, if the variables that theory predicts should explain investment in information play no role, one could argue that our indicator of information or the explanatory variables are fraught with errors. Third, the estimates of information investment might provide indirect evidence on overconfidence. If variables which tend to be associated with overconfidence - such as gender - have no effect on information, one may also doubt that overconfidence affects investors' decisions. Finally, estimates of information investment help identifying variables that can be used as instruments when, later in the paper, we estimate the effect of information on portfolio performance.

The rational model in Section 2 suggests that three variables should affect investment in information: wealth, risk tolerance and the marginal cost of collecting financial information. Figure 1 plots investment in financial information (measured in minutes per week) against financial wealth. The relation is strongly positive, particularly at low levels of wealth. Figure 2 shows that information is negatively correlated with the risk aversion indicator: investors who report to be risk averse invest much less in information than investors who are more risk tolerant.

Figure 3 plots information against education. We have no direct measure of the cost of information, and proxy it with years of schooling. Education reduces the cost of acquiring information because investors with higher education need less time to obtain an extra unit of information. On the other hand, information requires time, and since higher education is associated with higher wages, investors with higher education also face a higher marginal cost of time. In the regression analysis we use also a dummy for retirement as a proxy for the value of time, and our expectation is that retired investors spend more time in gathering financial information. Empirically, we find a positive association between education and information, consistent with the hypothesis that investors with higher education have a lower cost of information. Since education is also positively correlated with the value of time, the coefficient is a lower bound of the cost effect of education.[8]

The regression analysis in Table 4 confirms the two-way correlations. Given the categorical nature of the dependent variable, the estimates are performed by ordered probit. We use three dummies for risk aversion, excluding the most risk-averse group. Even when financial wealth, risk aversion and education are introduced simultaneously, each variable has an independent and statistically significant effect on investment in information. The economic impact of these variables, however, is rather different. Raising financial wealth from the bottom to the top quartile lowers the probability of making no information investment by only 2 percentage points (5% of the sample mean). Risk tolerance has a much stronger impact: being in the highest risk tolerance group lowers the probability of not acquiring information by 26 points (75% of the sample mean); increasing education by 5 years (one standard deviation) lowers the probability of no information investment by 9 points.

Overall, these correlations lend support to the loop first investigated by Arrow (1987) whereby individuals can increase their payoff by acquiring information on rates of return. Because the value of information is directly related to the amount to be invested, the wealthy have stronger incentive to acquire information, increasing the expected rate of return. It is this mechanism that makes the distribution of final wealth more unequal than that of initial wealth. This mechanism can be reinforced even further if the wealthy are also more risk tolerant and better educated, enhancing their incentive to obtain information and thus higher returns.

[8] An alternative interpretation is that those with higher education have a preference for finance. Some individuals may obtain utility from collecting financial information; for them the marginal benefit of financial information is even larger and thus they invest more in financial information. Even if these preferences are unobserved they will be reflected in the information acquired. Having raised this issue, note that unobserved taste for financial information does not necessarily affect the implications of the two models. If investors are rational, those who like finance purchase more information. But they also benefit more from information, and the Sharpe ratio is still positively correlated with information. If investors are overconfident, those who purchase more information for pleasure are also hurt more: information and the Sharpe ratio are negatively correlated, because investors are overconfident, not because they like finance. We come back to this issue in Section 5.

In column 2 of Table 4 we add a dummy for retirement as a further proxy of the cost of information, and an indicator of income risk. This indicator equals one if the respondent is unable to predict if his or her income will fall significantly, increase significantly or remain unchanged in the 5 years following the interview. In more general models, any variable - such as income risk - that affects the demand for stocks should also affect the demand for information. For instance, those who expect to allocate less wealth to stocks, e.g. because of high income risk, also benefit less from information. Consistent with this interpretation, income risk is negatively associated with information. The coefficient of the dummy for retirement is positive as expected.

Column 3 adds other demographic variables to account for variation in preferences which are possibly correlated with wealth, education or risk aversion: region (a dummy for living in the North), gender, marital status and city size. The results are qualitatively unchanged, suggesting that the correlations between financial information and wealth, education and risk aversion are not due to omitted demographic characteristics.

Controlling for gender is particularly important in the present context. Previous empirical literature suggests that men tend to be more overconfident than women in relation to male specific tasks, such as finance (Lundeberg et al., 1994; Barber and Odean, 2001). The positive coefficient of the male dummy is consistent with this evidence. The probability that males spend no time in information is 33 percentage points lower than females, while the probability of spending more than two hours per week is 45 percentage points higher. Of course, we cannot rule out that the male dummy reflects omitted variables correlated with gender.

The other regressions in Table 4 report various sensitivity checks. In column 4 we replace the dummies for risk aversion with the alternative measure based on the respondents' opinion about risk. Viewing risk as a threat rather than as an opportunity is negatively associated with investment in information, but the other results are unchanged.

Column 5 includes only stockholders, since acquiring information is mostly relevant for them and those who don't have stocks may provide inaccurate answers; results are again similar to the total sample estimates. Finally, column 6 drops those who spend more than 7 hours per week to make sure that the correlations between information and wealth, risk tolerance and education are not driven by a small group of outliers with above-average taste for financial information. The estimates are again unaffected.

Overall, the estimates are consistent with the hypothesis that those who invest in information do it because they expect, rightly (as in the rational model) or wrongly (as in the overconfidence model), to benefit from it not only in terms of average returns but also in terms of risk adjusted returns. In the next section we test whether, in fact, they are right or wrong.

5. Information and portfolio performance

The regressions for the Sharpe ratio in Table 5 represent the core estimates of the paper. Since the Sharpe ratio is not defined for individuals who don't have risky assets, we have valid observations for 1,365 investors. Of these, 80% have accounts only with Unicredit, while 20% also with other banks. In the latter case, we observe both wealth components.

5.1. Baseline results

Column 1 reports OLS estimates using the indicator of financial information as the only explanatory variable. In a model where investors are free from psychological bias, cross-sectional differences in the Sharpe ratio arise only from differences in correctly processed information. Contrary to the prediction of the rational model, the coefficient of information is negative and statistically different from zero at the 1 percent level. The effect is also quantitatively large: those who spend between 2 and 4 hours per week in information have a Sharpe ratio that is 27% lower than those who spend no time. Increasing time spent in information from 30 minutes per week (the median) to 2-4 hours (the 90th percentile) lowers the Sharpe ratio by 13.5%. At the sample median of the portfolio standard deviation, this is equivalent to a 17 basis points reduction in the portfolio expected excess return.

The estimates may be affected by selection bias because, as noted above, the Sharpe ratio is defined only for investors with positive amounts of stocks. And some may choose not to invest in the stock market precisely because they receive bad signals from the market.

To account for this source of selection bias, in column 2 we report the second stage regression of a Heckman two-step estimator. The first stage is a probit regression where the decision to invest in risky assets depends on investment in information, financial wealth (linear and quadratic terms), risk aversion and demographic variables. Identification is obtained omitting financial wealth from the second stage regression for the Sharpe ratio. The restriction is implied by the model of Section 2: if there are fixed transaction costs, financial wealth affects the decision to invest in risky assets, but it does not affect the Sharpe ratio once information is controlled for.[9] The results are similar: the coefficient of information is still negative and statistically different from zero, and its magnitude is only slightly reduced.

Column 3 of Table 5 adds dummies for region, gender, marital status and city size. In the rational model these variables should not affect the Sharpe ratio, unless they proxy for differences in information not captured by our indicator. The coefficients of these additional variables are jointly not statistically different from zero.

The results can be criticized for three reasons. First, the negative correlation between information and the Sharpe ratio may reflect unobserved factors (not captured by the demographic variables) that affect portfolio performance and are correlated with financial information. For instance, ability to manage the portfolio differs across investors, and smart investors could achieve a higher Sharpe ratio without spending too much time in collecting information. Time spent in information would then be negatively correlated with unobserved ability, resulting in a negative correlation with the Sharpe ratio. A second criticism is that the negative correlation may be the result of a systematic downward bias in measured returns resulting from unobserved taste for finance. Some

[9] The first stage results indicate that those who invest in information are more likely to invest in stocks. Causality however can run both ways depending on the timing of the participation decision and information acquisition. The coefficient of risk tolerance is positive. Wealth has a strong positive effect on participation, consistent with the presence of fixed transaction costs.

investors may trade and invest in risky assets because they like it, but the utility gain from the extra risk is not reflected in the monetary portfolio payoff. Furthermore, since these investors enjoy finance they also spend more time collecting information, hence the negative correlation. Finally, if the information variable is measured with error the estimates are biased towards zero.

These concerns imply that our information indicator might be correlated with the regression error, producing biased estimates. We address these concerns using an instrumental variable approach. We use as instruments the indicator of income risk and the retirement dummy. As shown in Table 4, both variables predict investment in information and there is no obvious reason why they should affect portfolio performance directly or be correlated with a taste for finance.

Column 4 of Table 5 reports the selectivity adjusted IV estimates. The coefficient of information is negative, precisely estimated, and larger in absolute terms than in the OLS estimates. The Sargan test of the over-identifying restrictions does not reject the null hypothesis that the instruments are orthogonal to the error term. The value of the F test for the excluded instruments in the first stage regression suggests that the estimates do not suffer from a weak instrument problem.

Column 5 adds to the second stage IV estimates three dummies for risk tolerance. In the rational model, risk tolerance should not affect the Sharpe ratio, once differences in information are controlled for. If our variables control imperfectly for differences in information, the correlation between risk tolerance and the Sharpe ratio should be positive, because risk tolerance and information are positively correlated, providing a supplementary test of the rational model. We find that risk tolerance is negatively correlated with portfolio performance: the Sharpe ratio of the most risk tolerant group is 7.8 percentage points lower than that of the least risk tolerant (the excluded category). This result contrasts with the rational model; to the extent that overconfidence is positively correlated with risk tolerance, it may be consistent with the overconfident model. The last regression in Table 5 excludes investors who spend more than 7 hours per week collecting information. The information coefficient is unaffected, implying that the results are not driven by a small group of investors with a taste for finance.

Table 6 repeats the OLS and IV estimates restricting the sample to investors with accounts only at Unicredit. For these investors the administrative data provide a complete coverage of the household portfolio which is not affected by measurement error. The sample selection results in a loss of 451 households with multiple bank relations, so the sample size drops to 914 observations. The pattern of the estimates is unaffected: the portfolio Sharpe ratio is negatively correlated with investors' information and the result is robust to selection and correction for unobserved heterogeneity. The notable difference with respect to the estimates in Table 5 is that some of the demographic variables (gender and residence, in particular), affect the Sharpe ratio.

The negative relation between the Sharpe ratio and information begs the question of why informed investors attain a lower Sharpe ratio. Is it because their returns are "too low" or because risk is "too high"? To distinguish between these two possibilities, Table 7 reports regressions relating the expected return and standard deviation of the portfolio to financial information. Investors who collect more information obtain higher returns (the coefficient is 0.135), consistent with Arrow (1987) conjecture and the prediction of portfolio models with endogenous information collection. However, the portfolio volatility is strongly increasing in information, driving the negative correlation between the Sharpe ratio and information reported in Table 5, in line with models of overconfident investors.

5.2. Sample splits by proxies of overconfidence

To assess the role of overconfidence, we exploit the theoretical implication that the negative correlation between information and the Sharpe ratio should be stronger for investors that, a priori, can be classified as "more overconfident", as in Figure 1. Experimental evidence shows that overconfidence differs considerably across individuals and tasks (West and Stanovich, 1997). When individuals are subject to multiple experiments over different domains, those who show more overconfidence in one domain - e.g. a classical knowledge-based test of overconfidence - tend to exhibit also more confidence in other domains. This suggests that there are traits that are specific to individuals (rather than to tasks) that affect the degree of overconfidence. There is indeed evidence that stable characteristics such as gender (e.g. Lundeberg et al., 1994; and Odean, 2001) or physical traits such as height (Addoum et al., 2017) predict overconfidence. One reasons is that in tasks that are specific to a type, individuals of that type exhibit more overconfidence. In particular, in more masculine tasks, such as finance, males show more overconfidence than women and vice versa. Experimental research also shows that overconfidence is more likely to manifest itself when individuals face relatively difficult tasks, such as finance (Fischhoff et al. 1977; and Yates (1990).

We split the sample using two different proxies for overconfidence. The first proxy is based on how well survey participants think they know stocks. One robust finding of the experimental literature is that when problems are grouped according to confidence level, the greatest overconfidence is observed for the problems answered with the greatest confidence, see Klayman et al. (1999). Furthermore, several studies suggest that overconfident individuals tend to overestimate their knowledge, see Weinstein (1980), Svenson (1981), and Taylor and Brown (1988). Accordingly, we classify as overconfident those who claim they know stocks well or very well (56 percent of the sample). The second split is based on gender, on the assumption that finance is typically a masculine task, as suggested by Barber and Odean (2001). In our sample males are responsible for financial matters of the household in 75 percent of the cases (85 percent excluding singles).

Results are reported in Table 8. The coefficient of information is more negative in the groups that are classified as more overconfident: males and those who claim to know stocks well. To provide a sense of the magnitudes involved, we compute the percent reduction in the Sharpe ratio when time spent increases from 30-60 minutes to 2-4 hours per week. Evaluated at sample means, the reduction in the ratio in the high overconfidence groups are between 10 and 20 percent higher. The results are consistent with the hypothesis that the propensity to take financial risk increases with overconfidence.

5.3. Robustness analysis using the 2007 sample

In the 2004 Unicredit Survey the Sharpe ratio is estimated using a simple cross section of portfolio weights and a times series of asset returns and covariances. One concern is that the weights change in response to information and this is not well captured by the portfolio composition of a single cross section of portfolios. To address this issue, we merge the 2007 Survey with administrative data with monthly information on 26 types of assets from January 2006 to September 2010.

To compute the Sharpe ratio, we first classify the 26 assets and in the same asset categories as in the 2004 survey: risk free, medium term government bonds, long term government bonds and stocks. We then define the weights of each category as the ratio of its corresponding value and the total value of the portfolio in each month. Expected returns, standard deviations and Sharpe ratios of the portfolio are calculated for each household and for each month using the same procedure as for the 2004 Survey. The Sharpe ratio used in the regressions is then the average Sharpe ratio for each household over the period January 2006 to September 2010 (excluding years after the crisis does not affect the results). Finally, we merge the household level Sharpe ratio computed as described from the administrative records with household level demographic variables, risk aversion and investment in financial information from the 2007 survey to obtain our final sample.

The regressions of Table 9 replicate, and confirm, the analysis of Table 4. Wealth, risk tolerance and education are positively associated with investment in information, and each of the coefficients is statistically different from zero. The magnitude of the wealth coefficient is lower in the 2007 Survey than in the 2004 Survey (0.27 against 0.62). One reason is that the wealth measure in 2007 is somewhat less accurate than in 2004 as it does not include financial wealth held in intermediaries other than Unicredit. Results are similar when we include other demographic variables (column 2) and trim the sample excluding investors who spend more than 76 hours per week acquiring financial information.

The regressions of Table 10 use as dependent variable the Sharpe ratio and are directly comparable with those of Table 5. Column (1) refers to the total sample of 928 investors. The coefficient of the Sharpe ratio is negative and statistically different from zero, but somewhat smaller than in Table 5. Splitting the sample by proxies of overconfidence confirms that the coefficient of information is more negative in the groups that we classify as more overconfident: those who claim to know stocks well (column 3) and males (column 4). Overall, also the results based on the 2007 Survey are qualitatively remarkably similar to those using 2004 data, lending support to the overconfidence model.

6. Conclusions

Investment in financial information differs considerably across investors. There is also a lot of heterogeneity in portfolio allocations, portfolio returns and volatility, raising naturally the question of what is the relation between financial information and portfolio performance. Models with rational investors recognize that information is valuable and that investors have different endowments and preferences. Accordingly, investors purchase different amounts of information, and those who purchase more information achieve higher returns more efficient portfolio allocations, as summarized by the portfolio Sharpe ratio. Therefore, in models with rational agents, investment in information and the Sharpe ratio are positively correlated.

This implication is not borne out in representative surveys of Italian investors. Investors gather information according to the predictions of portfolio models with endogenous investment in financial information. And more informed investors do indeed obtain higher portfolio returns, contributing to explain why returns to wealth differ systematically across investors, a mechanism relevant for the current debate on the sources of wealth inequality. But we find that investors that acquire more information attain lower returns per unit of risk (a lower Sharpe ratio). This is not due to selection bias or omitted variables, because the correlation is still negative and even stronger when we instrument our proxy for financial information acquisition and when we account for endogenous selection of stock market participants.

We argue that the empirical correlation between the portfolio Sharpe ratio and investors' information is more easily understood if one allows investors to be overconfident about the quality of their information, while retaining expected utility maximization. Overconfident, but otherwise rational investors, collect information responding to the same economic incentives as rational investors but, compared to the former, collect too much information and rely too much on it. For moderate amounts of overconfidence, the correlation between the amount of information and the Sharpe ratio is actually negative. Our findings are consistent with these predictions. Furthermore, the negative relation between the Sharpe ratio and information is stronger among investors that can be classified as more overconfident.

Overconfident investors realize lower Sharpe ratios but obtain higher portfolio returns. Their wealth will tend to be higher than otherwise but it will also be more volatile. In future research it will be interesting to investigate how these two effects of financial information – higher portfolio returns coupled with a more than proportional increase in volatility – affect wealth inequality.

References

Addoum J. M., G. Korniotis, and Kumar, A. (2017), "Stature, Obesity, and Portfolio Choice", *Management Science* 63, 3393-3413.

Arrow, K. J. (1987), "The Demand for Information and the Distribution of Income", *Probability in the Engineering and Informational Sciences* 1, 3–13.

Bach, Laurent, Laurent Calvet and Paolo Sodini (2017), "Rich Pickings? Risk, Return, and Skill in the Portfolios of the Wealthy," CEPR Discussion Paper No. 11734.

Barber, B., and T. Odean (1999), "Do Investors Trade Too Much?," *American Economic Review* 89, 1279-98.

Barber, B., and T. Odean (2001), "Boys Will Be Boys: Gender, Overconfidence, and Common Stock Investment," *Quarterly Journal of Economics* 116, 261-92.

Barberis, N., and R. Thaler (2003), "A Survey of Behavioral Finance," *Handbook of the Economics of Finance*, G.M. Constantinides, M. Harris and R. Stulz eds. Amsterdam: Elsevier Science.

Barlevy, G., and P. Veronesi (1999), "Information Acquisition in Financial Markets," *Review of Economic Studies* 67, 79-90.

Benhabib, J., and A. Bisin (2018), "Skewed Wealth Distributions: Theory and Empirics," *Journal of Economic Literature* (forthcoming).

Best, J., and K. Dogra (2017), "Capital Management and Inequality", University of Oxford, mimeo.

Biais, B., D. Hilton, K. Mazurier and S. Pouget (2005), "Judgmental Overconfidence, Self-Monitoring, and Trading Performance in an Experimental Financial Market," *The Review of Economic Studies* 72, 287-312.

Bilias, Y., D. Georgarakos and M. Haliassos (2010), "Portfolio Inertia and Stock Market Fluctuations," *Journal of Money, Credit and Banking* 42, 715-742.

Boyle, P., L. Garlappi, R. Uppal and T. Wang, (2012), "Keynes Meets Markowitz: The Trade-Off Between Familiarity and Diversification," *Management Science* 2, 253-272.

Calvet, L. E., J.Y. Campbell and P. Sodini (2007), "Down or Out: Assessing the Welfare Costs of Household Investment Mistakes," *Journal of Political Economy* 115, 707-47.

Calvet, L. E., J.Y. Campbell and P. Sodini (2009), "Fight or Flight? Portfolio Rebalancing by Individual Investors, "*The Quarterly Journal of Economics* 124, 301-348.

Dohmen, T.D., A. Falk, D. Huffman, J. Schupp, U. Sunde and G. Wagner (2011), "Individual Risk Attitudes: Measurement, Determinants and Behavioral Consequences", *Journal of the European Economic Association* 9, 522-550

Fischhoff, B., P. Slovic, and S. Lichtenstein (1977), "Knowing with Certainty: The Appropriateness of Extreme Confidence," *Journal of Experimental Psychology* 3, 552-64.

Grossman, S. J., and J.E. Stiglitz (1980), "On the Impossibility of Informationally Efficient Markets," *American Economic Review* 70, 393-408.

Guiso, L. and M. Paiella (2008), "Risk Aversion, Wealth, and Background Risk," *Journal of the European Economic Association* 6, 1109-50.

Fagereng, A., L. Guiso, D. Malacrino and L. Pistaferri (2018), "Heterogeneity and Persistence in Returns to Wealth," under review

Huberman, G. (2001), "Familiarity Breeds Investment," *Review of Financial Studies* 14, 659-680.

Ivkovic, Z., C. Sialm, and S. Weisbenner (2008), "Portfolio Concentration and the Performance of Individual Investors," *Journal of Financial and Quantitative* Analysis 43, 613-56.

Kacperczyk, M., J. B. Nosal, and L. Stevens (2018), "Investor Sophistication and Capital Income Inequality", CEPR Discussion Paper No. 12870.

Klayman, J., J. B. Soll, C. Gonzales-Vallejo, and S. Barlas (1999), "Overconfidence: It Depends on How, What, and Whom You Ask," *Organizational Behavior and Human Decision Processes* 79, 216-47.

Lundeberg, M. A., P.W. Fox, and J. Puncochar (1994), "Highly Confident but Wrong: Gender Differences and Similarities in Confidence Judgments," *Journal of Educational Psychology* 86, 114-21.

Massa, M., and A. Simonov (2006), "Hedging, Familiarity and Portfolio Choice," *Review of Financial Studies* 19, 633-685.

Mondria, J., (2010), "Portfolio Choice, Attention Allocation, and Price Comovement," *Journal of Economic Theory* 145, 1837–1864.

Odean, T. (1998a), "Volume, Volatility, Price and Profit when all Investors are Above Average," *Journal of Finance* 53, 1887-1934.

Odean, T. (1998b), "Are Investors Reluctant to Realize their Losses?," *Journal of Finance* 53, 1775-98.

Pelizzon, L., and G. Weber (2005), "Are Households' Portfolios Efficient? An Analysis Conditional on Housing," *Journal of Financial and Quantitative Analysis* 43, 401-31.

Peress, J. (2004), "Wealth, Information Acquisition and Portfolio Choice," *The Review of Financial Studies* 17, 879-914.

Svenson, O. (1981), "Are We All Less Risky and More Skillful than our Fellow Drivers?" *Acta Psychologica* 47, 143-8.

Taylor, S. E. and J. D. Brown (1988), “Illusion of Well-Being: A Social Psychological Perspective on Mental Health,” *Psychological Bulletin* 103, 193-210.

Verrecchia, R.E. (1982), “Information Acquisition in a Noisy Rational Expectations Economy,” *Econometrica* 50, 1415-30.

Yates, J. F. (1990), *Judgment and Decision Making*. Englewood Cliffs, NJ: Prentice Hall.

Weinstein, N. D. (1980), “Unrealistic Optimism about Future Life Events,” *Journal of Personality and Social Psychology* 39, 806-20.

Van Nieuwerburgh, S., and L. Veldkamp (2010), “Information Acquisition and Under-diversification,” *Review of Economic Studies* 77, 779–805.

West, R. F., and K.E. Stanovich (1997), “The Domain Specificity and Generality of Overconfidence: Individual Differences in Performance Estimation Bias,” *Psychonomic Bulletin and Review* 4, 387-92.

Figure 1
Investment in information and financial wealth

Figure 2
Investment in information and risk aversion

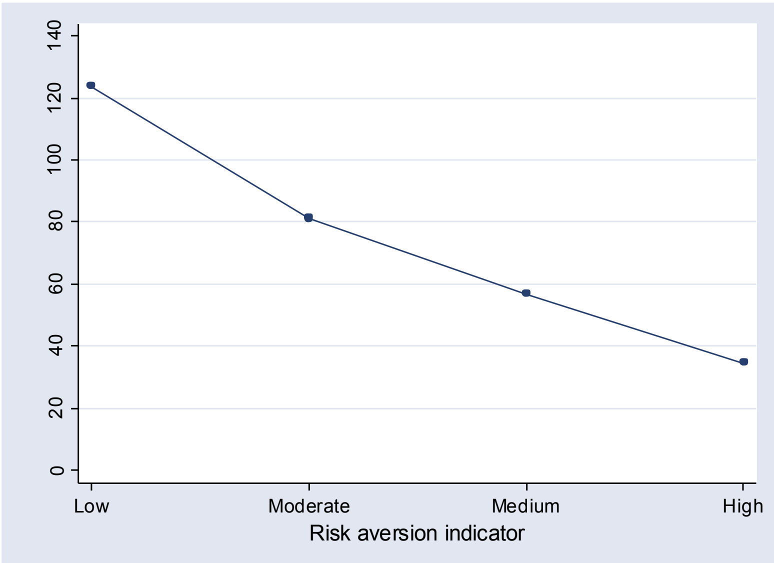


Figure 3
Investment in information and education

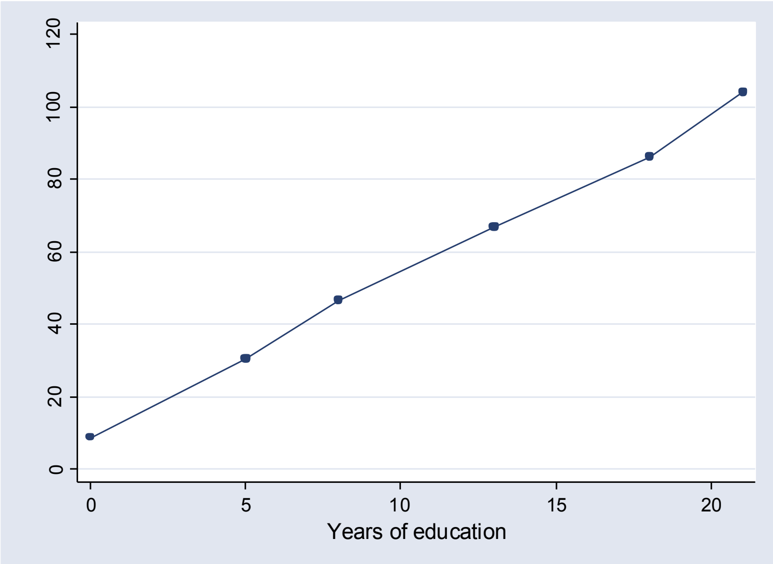


Table 1
Effect of information and risk tolerance on portfolio performance

	Model with rational investors		*Model with overconfident investors*	
	Effect of information	Effect of risk tolerance	Effect of information	Effect of risk tolerance
Portfolio expected return	+	+	+	+
Portfolio standard deviation	+	+	+	+
Sharpe ratio	+	0	- (more negative if more overconfident)	- (if risk tolerance is correlated with overconfidence)

Table 2
Investment in financial information

Time spent collecting financial information	No time	Less than 30 minutes	30-60 minutes	1 to 2 hours	2 to 4 hours	4 to 7 hours	More than 7 hours
% of investors	36.5	24.8	14.7	10.9	6.5	2.8	3.6
Equivalent number of working days in a year	0	1.5	4.5	8.4	18	33	42
% owning stocks	59.2	82.0	85.2	95.0	93.3	98.1	98.5
% invested in stocks	12.6	21.8	24.2	31.0	35.6	38.0	43.1

Note. The table reports the sample distribution of time spent in financial information in a typical week.

Table 3
Summary statistics

	Mean	Standard deviation
Investment in information		
Time spent collecting financial information	2.09	1.36
Financial wealth and portfolio performance		
Respondent's financial wealth ('000 euro)	40.0	170.3
Household's financial wealth ('euro)	90.6	375.4
Expected return of the portfolio	1.02	0.44
Standard deviation of the portfolio	3.69	4.73
Sharpe ratio	0.27	0.15
Share of risky assets in mutual funds (portfolio diversification)	0.58	0.44
Risk aversion, trading and delegation		
Low risk aversion	0.02	0.15
Moderate risk aversion	0.25	0.44
Medium risk aversion	0.47	0.50
High risk aversion	0.25	0.43
Risk is an opportunity	0.26	0.44
Trading activity (trades per month)	0.23	1.14
Demographic variables		
Age	51.7	15.0
Male	0.68	0.46
Married	0.65	0.47
Living in the North	0.75	0.43
Living in a city	0.51	0.49
Years of education	11.1	4.23

Note. The table reports summary statistics for the variables used in the estimation. Means and standard deviations are computed using population weights. See Appendix B for variables' definitions.

Table 4

Determinants of investment in financial information

	Total sample				*Stockholders only*	*Trimmed sample*
	(1)	(2)	(3)	(4)	(5)	(6)
Financial wealth	0.619 (0.092)**	0.548 (0.093)**	0.478 (0.094)**	0.487 (0.094)**	0.334 (0.095)**	0.454 (0.099)**
Years of education	0.049 (0.006)**	0.056 (0.006)**	0.060 (0.006)**	0.065 (0.006)**	0.052 (0.007)**	0.056 (0.007)**
Retired		0.270 (0.053)**	0.221 (0.054)**	0.189 (0.053)**	0.109 (0.060)	0.180 (0.055)**
Low risk aversion	0.919 (0.147)**	0.983 (0.148)**	0.972 (0.148)**		0.917 (0.165)**	0.879 (0.157)**
Moderate risk aversion	0.561 (0.076)**	0.588 (0.076)**	0.567 (0.076)**		0.449 (0.087)**	0.514 (0.078)**
Medium risk aversion	0.356 (0.072)**	0.374 (0.072)**	0.367 (0.072)**		0.285 (0.083)**	0.381 (0.073)**
Income risk		-0.161 (0.059)**	-0.154 (0.059)**	-0.161 (0.059)**	-0.131 (0.066)*	-0.124 (0.060)*
Risk is an opportunity				0.152 (0.056)**		
Male			0.437 (0.061)**	0.451 (0.060)**	0.468 (0.068)**	0.412 (0.061)**
Married			0.086 (0.058)	0.088 (0.058)	0.086 (0.065)	0.078 (0.059)
Resident in the North			0.325 (0.053)**	0.314 (0.052)**	0.274 (0.059)**	0.343 (0.054)**
Resident in a small city			-0.038 (0.053)	-0.042 (0.053)	-0.012 (0.060)	-0.009 (0.054)
Observations	1,834	1,834	1,834	1,834	1,419	1,767

Note. Ordered probit estimates for time spent to acquire financial information. The trimmed sample excludes investors who spend more than 7 hours per week. Standard errors are reported in parenthesis. Two stars denote significance at 1% or less; one star significance at 5% or less.

Table 5
Sharpe ratio and investment in financial information: total sample

	OLS	Selection adjusted		IV-Selection Adjusted		
	(1)	(2)	(3)	(4)	(5)	(6)
Investment in information	-0.018	-0.017	-0.016	-0.079	-0.052	-0.057
	(0.002)**	(0.003)**	(0.003)**	(0.021)**	(0.020)**	(0.020)**
Male			-0.015	0.019	0.008	0.004
			(0.010)	(0.016)	(0.015)	(0.013)
Married			-0.007	-0.013	-0.011	-0.014
			(0.009)	(0.011)	(0.010)	(0.010)
Resident in the North			-0.003	-0.015	-0.016	-0.015
			(0.008)	(0.010)	(0.009)	(0.009)
Resident in a small city			0.003	0.006	0.006	0.010
			(0.008)	(0.009)	(0.009)	(0.009)
Low risk aversion					-0.078	-0.102
					(0.029)**	(0.028)**
Moderate risk aversion					-0.078	-0.082
					(0.013)**	(0.013)**
Medium risk aversion					-0.046	-0.045
					(0.012)**	(0.013)**
Mills ratio		0.006	0.003	-0.217	-0.149	-0.132
		(0.017)	(0.018)	(0.076)**	(0.071)*	(0.054)*
Sargan test				1.311	0.876	1.065
p-value				(0.252)	(0.349)	(0.302)
F-test for excluded instruments				16.01	13.73	23.04
Observations	1,365	1,780	1,780	1,780	1,780	1,780

Note. The dependent variable is the Sharpe ratio, computed as the ratio of the portfolio expected excess return and the portfolio standard deviation. Column 1 reports OLS estimates, the other columns the second stage estimates of a Heckman selection model. The IV-Selection adjusted estimates use as instruments dummies for income risk and retirement. The sample includes only those with financial investment. The last column excludes investors who spend more than 7 hours per week in information. Standard errors are reported in parenthesis. Two stars denote significance at 1% or less; one star significance at 5% or less.

Table 6
Sharpe ratio and investment in financial information:
sample of investors with only one bank relation

	OLS	Selection adjusted		IV-Selection Adjusted		
	(1)	(2)	(3)	(4)	(5)	(6)
Investment in information	-0.011	-0.018	-0.018	-0.033	-0.027	-0.029
	(0.003)**	(0.003)**	(0.003)**	(0.010)**	(0.010)**	(0.012)*
Male			-0.005	0.004	0.001	0.005
			(0.009)	(0.011)	(0.011)	(0.011)
Married			-0.025	-0.028	-0.028	-0.031
			(0.009)**	(0.009)**	(0.009)**	(0.009)**
Resident in the North			-0.015	-0.019	-0.019	-0.022
			(0.008)	(0.009)*	(0.009)*	(0.009)*
Resident in a small city			0.002	0.004	0.003	0.004
			(0.008)	(0.008)	(0.008)	(0.008)
Low risk aversion					-0.037	-0.039
					(0.026)	(0.027)
Moderate risk aversion					-0.036	-0.036
					(0.012)**	(0.012)**
Medium risk aversion					-0.031	-0.033
					(0.011)**	(0.011)**
Mills ratio		-0.104	-0.122	-0.170	-0.162	-0.167
		(0.019)**	(0.020)**	(0.036)**	(0.036)**	(0.034)**
Sargan test				0.001	0.024	0.000
p-value				(0.980)	(0.877)	(0.996)
F-test for excluded instruments				18.04	15.73	23.04
Observations	914	914	914	914	914	868

Note. The dependent variable is the Sharpe ratio, computed as the ratio of the portfolio expected excess return and the portfolio standard deviation. The sample is restricted to households that have accounts with only one bank. Column 1 reports OLS estimates, the other columns the second stage estimates of a Heckman selection model. The IV-Selection adjusted estimates use as instruments dummies for income risk and retirement. The sample includes only those with financial investment. The last column excludes investors who spend more than 7 hours per week in information. Standard errors are reported in parenthesis. Two stars denote significance at 1% or less; one star significance at 5% or less.

Table 7
Financial information, excess return and standard deviation of the portfolio

	Excess return			Standard deviation		
	(1)	(2)	(3)	(4)	(5)	(6)
Investment in information	0.135 (0.008)**	0.127 (0.008)**	0.115 (0.008)**	0.999 (0.068)**	0.938 (0.070)**	0.823 (0.070)**
Male		-0.006 (0.031)	-0.011 (0.031)		0.290 (0.261)	0.238 (0.258)
Married		0.031 (0.030)	0.031 (0.030)		0.191 (0.252)	0.195 (0.248)
Resident in the North		0.171 (0.027)**	0.182 (0.027)**		0.736 (0.226)**	0.846 (0.223)**
Resident in a small city		-0.094 (0.027)**	-0.093 (0.027)**		-0.669 (0.225)**	-0.659 (0.221)**
Low risk aversion			0.304 (0.079)**			3.616 (0.653)**
Moderate risk aversion			0.221 (0.039)**			2.045 (0.320)**
Medium risk aversion			0.141 (0.036)**			1.182 (0.299)**
Observations	1,780	1,780	1,780	1,780	1,780	1,780

Note. OLS estimates of the relation between the portfolio expected return (columns 1-3) and standard deviation (columns 4-6) and investment in financial information. Standard errors are reported in parenthesis. Two stars denote significance at 1% or less; one star significance at 5% or less.

Table 8
Sharpe ratio and investment in financial information.
Sample splits by overconfidence indicators

	Low knowledge of stocks	High knowledge of stocks	Women	Men
	(1)	(2)	(3)	(4)
Investment in information	-0.006	-0.013	-0.009	-0.016
	(0.007)	(0.003)**	(0.006)	(0.003)**
Male	-0.025	-0.003		
	(0.016)	(0.012)		
Married	0.012	-0.015	-0.008	-0.003
	(0.016)	(0.011)	(0.016)	(0.011)
Resident in the North	-0.012	0.006	-0.038	0.011
	(0.016)	(0.010)	(0.017)*	(0.010)
Resident in a small city	-0.004	0.003	0.013	-0.000
	(0.015)	(0.010)	(0.016)	(0.009)
Mills ratio	0.014	0.012	0.041	0.007
	(0.029)	(0.026)	(0.029)	(0.022)
Observations	482	883	376	989

Note. Selectivity adjusted estimates of the relation between investment in information and the Sharpe ratio for various sample splits. The first stage probit of the two-stage Heckman estimator includes investment in information, financial wealth linear and square, three dummies for risk tolerance, education and demographics. Low and high knowledge of stocks split the sample between those who report knowing very well or well stocks, and those who don't. The sample includes only people with financial investment. Standard errors are reported in parenthesis. Two stars denote significance at 1% or less; one star significance at 5% or less.

Table 9

Determinants of investment in financial information - 2007 sample

	Total sample	Total sample	Trimmed sample
	(1)	(2)	(3)
Financial wealth	0.271***	0.274***	0.286***
	(0.051)	(0.051)	(0.054)
Education	0.023***	0.024***	0.017**
	(0.008)	(0.008)	(0.008)
Retired	0.063	0.018	-0.022
	(0.070)	(0.071)	(0.073)
Low risk aversion	1.229***	1.083***	0.881***
	(0.220)	(0.221)	(0.240)
Moderate risk aversion	0.659***	0.563***	0.520***
	(0.111)	(0.113)	(0.115)
Medium risk aversion	0.298***	0.241**	0.240**
	(0.104)	(0.105)	(0.106)
Male		0.658***	0.588***
		(0.076)	(0.077)
Married		0.041	0.062
		(0.072)	(0.074)
Resident in the North		0.004	0.063
		(0.065)	(0.067)
Resident in a small city		0.201	0.241
		(0.274)	(0.274)
Observations	1205	1205	1156

Note. Ordered probit estimates for time spent to acquire financial information using the 2007 survey. The trimmed sample excludes investors who spend more than 7 hours per week. Standard errors are reported in parenthesis. Two stars denote significance at 1% or less; one star significance at 5% or less.

Table 10
Sharpe ratio and investment in financial information - 2007 sample

	Total sample	Low knowledge of stocks	High knowledge of stocks	Women	Men
	(1)	(2)	(3)	(4)	(5)
Investment in information	-0.007**	0.004	-0.011***	0.008	-0.010***
	(0.003)	(0.006)	(0.003)	(0.007)	(0.003)
Male	-0.010	-0.026*	-0.003		
	(0.009)	(0.015)	(0.012)		
Married	0.018**	0.030**	0.014	0.009	0.021*
	(0.008)	(0.014)	(0.011)	(0.015)	(0.010)
Resident in the North	0.005	0.037***	-0.011	0.027*	-0.003
	(0.008)	(0.013)	(0.009)	(0.015)	(0.009)
Resident in a small city	0.047	0.062	-0.052	0.038	0.050
	(0.035)	(0.039)	(0.070)	(0.052)	(0.048)
Low risk aversion	0.044	0.038	0.052	-0.100	0.075**
	(0.029)	(0.053)	(0.036)	(0.074)	(0.031)
Moderate risk aversion	0.025*	-0.005	0.046**	0.020	0.031**
	(0.013)	(0.019)	(0.018)	(0.024)	(0.016)
Medium risk aversion	0.0094	0.003	0.020	-0.007	0.019
	(0.012)	(0.017)	(0.017)	(0.021)	(0.015)
Constant					
	0.191***	0.165***	0.196***	0.170***	0.183***
	(0.014)	(0.021)	(0.021)	(0.024)	(0.018)
Observations					
R-squared	928	314	614	272	656

Note. OLS estimates of the relation between investment in information and the Sharpe ratio for total sample and various sample splits using the 2007 survey. Low and high knowledge of stocks split the sample between those who report knowing very well or well stocks, and those who don't. The sample includes only people with financial investment. Standard errors are reported in parenthesis. Two stars denote significance at 1% or less; one star significance at 5% or less.

Appendix

June 13, 2018

A: The rational and the overconfident models

For our purpose, Peress (2004) framework is the most appropriate. As in Peress (2004), we consider a rational investor who chooses between a risk free asset (a "bond") and a risky asset (a "stock"). We first recall the expressions for the equilibrium price P, the optimal portfolio share of stocks α_j^R and the optimal level of information acquisition x_j^R. The equilibrium price, for small levels of z, which scales the level of risk in the economy is :

$$\ln P \approx pz = p_0(i)z + p_\pi(i)(\pi z - \mu\theta z) - r^f z$$

where:

$$\begin{aligned}
\pi z &\sim N((E\pi)z, \sigma_\pi^2 z) \text{ is the payoff of the risky asset} \\
\theta z &\sim N((E\theta)z, \sigma_\theta^2 z) \text{ is the supply of the risky asset} \\
p_0(i) &= \frac{1}{\bar{h}}\left(\frac{E\pi}{\sigma_\pi^2} + i\frac{E\theta}{\sigma_\theta^2} + \frac{1}{2}\right) \\
p_\pi(i) &= 1 - \frac{1}{\bar{h}\sigma_\pi^2} \\
\bar{h} &= \frac{1}{\sigma_\pi^2} + \frac{i^2}{\sigma_\theta^2} + \frac{i}{n} \text{ (aggregate precision)} \\
i &= \int_j x_j^R \tau_j(W_{0j}) dG(x_j^R, W_{0j}) \text{ (aggregate information)} \\
n &= \int_j \tau_j(W_{0j}) dG(x_j^R, W_{0j}) \text{ (aggregate risk tolerance)} \\
&\quad \tau_j(W_{0j}) \text{ (absolute risk tolerance)} \\
&\quad dG(x_j^R, W_{0j}) \text{ (density of investors with } x_j^R, W_{0j})
\end{aligned}$$

x_j^R is the amount of information acquired by investor j and is given by:

$$C'(x_j^R) = \frac{1}{2}\tau(W_j)\phi'_x(x_j^R, i)$$

where:

$$\phi(x_j^R, i) = h(i, x_j^R)A + \frac{1}{4h(i, x_j^R)} + q - 1 \text{ is increasing and convex in } x$$

$$
\begin{aligned}
A &= \frac{h(i,0)n^2 + 2in + \sigma_\theta^2}{(n\bar{h})^2} + q^2 \\
q &= \frac{E\theta}{n\bar{h}} - \frac{1}{2\bar{h}} \\
h(i, x_j^R) &= \frac{1}{\sigma_\pi^2} + \frac{i^2}{\sigma_\theta^2} + x_j^R
\end{aligned}
$$

Theorem 1 in Peress (2004) proves that the optimal amount of information is increasing in absolute risk tolerance $\tau(W_j)$ and wealth, and decreasing in the marginal cost of information. It also shows that there is a wealth threshold below which the investor does not purchase information.

The optimal share of stocks is:

$$
\alpha_j^R = \frac{\tau(W_j)}{W_j}\left(\frac{E\pi}{\sigma_\pi^2} + \frac{iE\theta}{\sigma_\theta^2} + \frac{i^2}{\sigma_\theta^2}(\pi - \mu\theta) + x_j^R \frac{S_j^R}{z} + \frac{1}{2} - (p + r^f)h(i, x_j^R)\right)
$$

where S_j^R is the signal received on the true payoff Π by investor j as he purchases the amount of information x_j^R :

$$
\begin{aligned}
S_j^R &= S(x_j^R) = \pi z + \sqrt{\frac{z}{x_j^R}}\varepsilon \\
\varepsilon &\sim N(0,1)
\end{aligned}
$$

0.1 The Sharpe ratio of the rational investor

The Sharpe ratio of the rational investor (dropping the index j for simplicity) is:

$$
Sharpe^R = \frac{E(\alpha^R r^e)}{\sqrt{V(\alpha^R r^e)}} = \frac{\text{portfolio mean excess return}}{\text{portfolio standard deviation}}
$$

where r^e is the excess return on the stock:

$$
r^e = \frac{\Pi - P}{P} - r^f z
$$

and the expected mean excess return on the portfolio is:

$$
\begin{aligned}
E(\alpha^R r^e) &= E(E(\alpha^R r^e |\ S^R, P)) \\
&= E\alpha^R E(r^e |\ S^R, P)
\end{aligned}
$$

Let us define

$$
\begin{aligned}
\lambda^R &= \alpha^R \frac{W}{\tau(W)} \frac{1}{\sqrt{V}} \\
&\quad \text{(the Sharpe ratio of the portfolio return)} \\
V &= V(\pi z \mid S^R, P)
\end{aligned}
$$

One can show that:

$$E(r^e|\ S^R, P) \simeq \lambda^R \frac{1}{\sqrt{V}}$$

So that:

$$\begin{aligned}
E(\alpha^R r^e) &= E(\left(\lambda^R\right)^2)\frac{\tau(W)}{W} \\
&= \frac{\tau(W)}{W}\phi(i, x^R)z
\end{aligned}$$

And the variance of the excess return on the portfolio is:

$$V(\alpha^R r^e) = E(V(\alpha^R r^e \mid S^R, P)) + V(E(\alpha^R r^e|\ S^R, P))$$

Since the mean excess return on the portfolio is of the order of z, $V(E(\alpha^R r^e|$ $S^R, P))$ is of the order of z^2 and is negligible at the first order in z with respect to $E(V(\alpha^R r^e \mid S^R, P))$. This implies that:

$$\begin{aligned}
V(\alpha^R r^e) &\simeq E(V(\alpha^R r^e \mid S^R, P)) \\
&= E(\left(\alpha^R\right)^2 V(r^e \mid S^R, P)) \\
&\simeq E(\left(\alpha^R\right)^2 V(\pi z|\ S^R, P)) \\
&= E(\left(\lambda^R\right)^2 \frac{1}{V} V \left(\frac{\tau(W)}{W}\right)^2 \\
&= \left(\frac{\tau(W)}{W}\right)^2 \phi(i, x^R)z
\end{aligned}$$

and finally the Sharpe ratio is given by:

$$Sharpe^R = \sqrt{\phi(i, x^R)z}$$

As shown in Peress (2004), Appendix B, the Sharpe ratio of the rational investor increases with x^R and the amount of information purchased.

0.2 The effect of overconfidence

We now introduce overconfidence in the previous model and compute the optimal portfolio and the optimal amount of information purchased by the overconfident investor. We assume that overconfident investors have mass zero among all other rational agents, so they don't affect the equilibrium price or the choice of other rational agents.

An overconfident investor who purchases the amount of information x^K thinks he is receiving the signal $S^K = \pi z + \sqrt{\frac{z}{Kx^K}}\varepsilon$ although he is actually receiving the signal $S(x^K)$. That is, he overestimates the true precision of the signal by a factor $K > 1$ measuring the degree of overconfidence. This alters the signal extraction problem he solves when he computes the optimal portfolio and chooses the optimal amount of information. Since the overconfident investor behaves as a rational investor who gets a signal with precision Kx^K,the optimal amount of information purchased is:

$$C'(x_j^K) = \frac{1}{2}\tau(W_j)K\phi_x'(Kx_j^K, i)$$

Following the same line of proof as Peress (2004), optimal information of the overconfident investor increases with risk tolerance, wealth and degree of overconfidence K, and decreases with the marginal cost of acquiring information. Here too, there is a threshold level of wealth below which the overconfident investor does not acquire information but, *ceteris paribus*, the threshold is lower than for the rational investor. This can be seen by noticing that information is not acquired if $C'(0) > \frac{1}{2}\tau(W_j)K\phi_x'(0)$, which requires a lower value of wealth the larger is K.

The optimal portfolio is given by:

$$\alpha_j^K = \frac{\tau(W_j)}{W_j}\left(\frac{E\pi}{\sigma_\pi^2} + \frac{iE\theta}{\sigma_\theta^2} + \frac{i^2}{\sigma_\theta^2}(\pi - \mu\theta) + Kx_j^K\frac{S^K}{z} + \frac{1}{2} - (p + r^f)h(i, Kx_j^K)\right)$$

Proof: The proof follows Peress (2004) except for the signal extraction problem. Now:

$$\begin{aligned} E(\pi z \mid S^K, P) &= \frac{1}{h(i, Kx_j^K)}\left(\frac{E\pi}{\sigma_\pi^2}z + \frac{iE\theta}{\sigma_\theta^2}z + z\frac{i^2}{\sigma_\theta^2}(\pi - \frac{\theta}{i}) + Kx_j^K S_j^K\right) \\ V(\pi z \mid S^K, P) &= \frac{z}{h(i, Kx_j^K)} \end{aligned}$$

and for small z the optimal portfolio is still given by :

$$\alpha_j^K = \frac{\tau(W_j)}{W_j}\frac{E(\pi z \mid S_j^K, P) + \frac{1}{2}V(\pi z \mid S_j^K, P) - pz - r^f z}{V(\pi z \mid S_j^K, P)}.$$

Substituting for the expected stock return and variance conditional on the signal gives the optimal portfolio choice above.

0.3 The perceived Sharpe ratio of an overconfident investor

The perceived Sharpe ratio of an overconfident investor can be computed exactly along the same lines followed for the rational investor but conditioning on the perceived signal. This gives:

$$Perceived\ Sharpe^K = \sqrt{\phi(i, Kx^K)z}$$

The perceived ratio increases with K (the overconfidence parameter) and x^K (the amount of information purchased). Thus, the overconfident investor is indeed tempted to purchase even more information than the rational investor. However, the true Sharpe ratio of the overconfident investor is different from the one he expects.

0.4 The actual Sharpe ratio of an overconfident investor

To compute the true Sharpe ratio of the overconfident investor, we need to condition on the true signal:

$$E(\alpha^K r^e) = E(E(\alpha^K r^e | \; S^R, P))$$

Define:

$$\begin{aligned}
\alpha^{RK} &= \alpha^K - \alpha^R \\
\lambda^{RK} &= \alpha^{RK} \frac{W}{\tau(W)} \sqrt{V}
\end{aligned}$$

where α^R denotes the portfolio choice of a rational investor who would be choosing the amount of information x^K.

Thus the average portfolio return of the overconfident investor, conditioning on the true distribution of signals, is:

$$\begin{aligned}
E(E(\alpha^K r^e | \; S^R, P)) &= E((\alpha^R + \alpha^{RK}) E(r^e | \; S^R, P)) \\
&= \frac{\tau(W)}{W} E(\lambda^R + \lambda^{RK}) \lambda^R \\
&= \frac{\tau(W)}{W} \left(E((\lambda^R)^2) + E(\lambda^{RK} \lambda^R) \right)
\end{aligned}$$

and the portfolio variance:

$$\begin{aligned}
V(\alpha^K r^e) &\simeq E(V(\alpha^K r^e \mid S^R, P)) \\
&= E(\alpha^{RK} + \alpha^R)^2 V(r^e \mid S^R, P) \\
&= \left(\frac{\tau(W)}{W}\right)^2 (E(\lambda^{RK} + \lambda^R)^2) \\
&= \left(\frac{\tau(W)}{W}\right)^2 (E(\lambda^{RK})^2 + E(\lambda^R)^2 + 2E(\lambda^R \lambda^{RK}))
\end{aligned}$$

Thus, the true Sharpe ratio of the overconfident investor is:

$$Sharpe^K = \frac{E((\lambda^R)^2) + E(\lambda^{RK} \lambda^R)}{\sqrt{E(\lambda^R)^2 + E(\lambda^{RK})^2 + 2E(\lambda^R \lambda^{RK})}}$$

Note that $E\left(\lambda^R\right)^2$ is the square of the Sharpe ratio of an investor who is not overconfident and perceives the signal correctly but who would purchase the amount of information x^K. We are interested in showing how the Sharpe ratio of the overconfident investor compares to that of the rational investor and how it varies with the amount of information x^K.

Let us compute the different terms that come into play:

$$\lambda^R = \frac{1}{\sqrt{zh(i, x^K)}} \times$$

$$\lambda^{RK} = \begin{array}{l} (p_0(i)(i/n - x^K)z + \pi z(1-p_\pi(i))(x^K - i/n) + \\ (\theta z/i)((1-p_\pi(i))i/n + p_\pi(i)x^K) + \sqrt{zx^K}\varepsilon) \\ \dfrac{1}{\sqrt{zh(i,x^K)}} \times \\ (-(K-1)xp_0(i)z + (1-p_\pi(i))\,(\pi z)\,(K-1)\,x^K + \\ p_\pi(i)(K-1)x^K(\theta z/i) + \varepsilon(\sqrt{K}-1)\sqrt{zx^K}) \end{array}$$

where ε is a random variable with a standard normal distribution (mean 0 and variance 1)

And where (to a first-order approximation in z):

$$E((\lambda^R))^2 = \phi(i, x^K)$$

$$\begin{aligned} E(\lambda^R \lambda^{RK}) &= \frac{x^K}{h(i,x^K)} \times \\ &\quad (((1-p_\pi(i))^2\sigma_\pi^2 + (p_\pi(i))^2\,\frac{\sigma_\theta^2}{i^2})(x^K - i/n)(K-1) + \\ &\quad p_\pi(i)\frac{i}{n}\frac{\sigma_\theta^2}{i^2}(K-1) + (\sqrt{K}-1)) \end{aligned}$$

and

$$\begin{aligned} E(\lambda^{RK})^2 &= \frac{\left(x^K\right)^2}{h(i,x^K)}((1-p_\pi(i))^2\sigma_\pi^2 + p_\pi(i)^2\sigma_\theta^2/i^2)(K-1)^2 \\ &\quad +(\sqrt{K}-1)^2\frac{x^K}{h(i,x^K)} \end{aligned}$$

Going back to the expression for the "true" Sharpe ratio of the overconfident investor, we have that:

$$Sharpe^K = \frac{N}{\sqrt{D}}$$

where:

$$\begin{aligned} N &= E(\lambda^R)^2 + E(\lambda^R\lambda^{RK}) \\ D &= E(\lambda^R)^2 + 2E(\lambda^R\lambda^{RK}) + E(\lambda^{RK})^2 \end{aligned}$$

Although it is not obvious from the formulas that $Sharpe^K < Sharpe^R$, we know that the amount of information purchased by the overconfident investor x^K is strictly greater than that purchased by the rational investor, so that the portfolio allocation of the overconfident investor is suboptimal given the equilibrium returns, implying $Sharpe^K < Sharpe^R$. (In equilibrium, the Sharpe ratio is maximized at the optimal level of information given the true signalling structure).

To see how the Sharpe ratio varies with a marginal increase in information , we compute the derivatives of the various terms of the ratio with respect to x^K :

$$\begin{aligned}
\frac{\partial E(\lambda^R)^2}{\partial x^K} &= \frac{\partial \phi(i,x^K)}{\partial x^K} = A - \frac{1}{4\left(h(i,x^K)\right)^2} \\
\frac{\partial E(\lambda^R\lambda^{RK})}{\partial x^K} &= (K-1)C\frac{x^K}{h(i,x^K)} + \\
& \quad (C(x^K - i/n)(K-1) + p_\pi(i)\frac{i}{n}\frac{\sigma_\theta^2}{i^2}(K-1) + (\sqrt{K}-1)) \\
& \quad \times \left(\frac{1}{h(i,x^K)} - \frac{x^K}{\left(h(i,x^K)\right)^2}\right) \\
\text{where } C &= (1-p_\pi(i))^2\sigma_\pi^2 + p_\pi(i)^2\sigma_\theta^2/i^2 \\
\frac{\partial E(\lambda^{RK})^2}{\partial x^K} &= \left(\frac{2x^K}{h(i,x^K)} - \frac{\left(x^K\right)^2}{\left(h(i,x^K)\right)^2}\right)C(K-1)^2 \\
& \quad +(\sqrt{K}-1)^2\left(\frac{1}{h(i,x^K)} - \frac{x^K}{\left(h(i,x^K)\right)^2}\right)
\end{aligned}$$

So that:

$$\begin{aligned}
\frac{\partial N}{\partial x^K} &= \frac{\partial E(\lambda^R)^2}{\partial x^K} + \frac{\partial E(\lambda^R\lambda^{RK})}{\partial x^K} > 0 \\
\frac{\partial D}{\partial x^K} &= \frac{\partial E(\lambda^R)^2}{\partial x^K} + 2\frac{\partial E(\lambda^R\lambda^{RK})}{\partial x^K} + \frac{\partial E(\lambda^{RK})^2}{\partial x^K} > 0
\end{aligned}$$

And finally:

$$\begin{aligned}
\frac{\partial Sharpe^K}{\partial x^K} &= \frac{1}{D}\left(\frac{\partial N}{\partial x^K}\sqrt{D} - N\frac{\frac{\partial D}{\partial x^K}}{2\sqrt{D}}\right) \\
&= \frac{1}{2D\sqrt{D}}\left(2\frac{\partial N}{\partial x^K}D - N\frac{\partial D}{\partial x^K}\right)
\end{aligned}$$

A marginal increase in the amount of information of overconfident investors has two effects. It increases the true excess mean return of the portfolio $\left(\frac{\partial N}{\partial x^K} > 0\right)$, but it also increases the true variance of the excess return $\left(\frac{\partial D}{\partial x^K} > 0\right)$.

Proof of Propositions 1 and 2. In general, it is not possible to establish analytically the sign of the above derivative. But we see that for large K, the extra term $E(\lambda^{RK})^2$ in the denominator of the Sharpe ratio of the overconfident investor dominates and the Sharpe ratio becomes:

$$Sharpe^K \sim B\frac{1}{h(i,x^K)} \text{ for large } K$$

where B is a positive constant. Since $h(i, x^K)$ is strictly increasing in x^K, when overconfidence is sufficiently large the Sharpe ratio is decreasing in K.

To see how the Sharpe ratio varies with the amount of information and the degree of overconfidence, we evaluate the ratio using the same assumptions as in Peress (2004, Section 6). In particular, we use a CRRA specification for the utility function with a coefficient of relative risk aversion of 5 to compute the level of aggregate risk tolerance n (using the same number for aggregate financial wealth of 5,184 billion dollars):

$$n = 1,037 \text{ billion (USD)}$$

and:

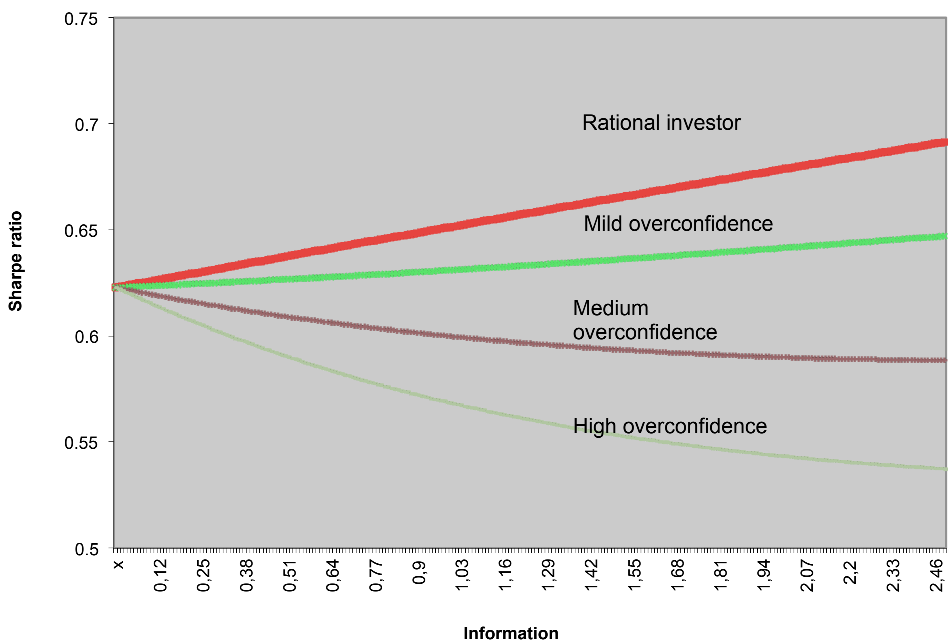
$$\begin{aligned}
\sigma_\pi^2 &= 0.0275 \text{ (the historical moments of stock returns in the US)} \\
E\theta &= n \times 2.750 \\
\sigma_\theta^2 &= n \times 6.539 \\
\mu &= 100 \times n
\end{aligned}$$

The computation is not meant to be realistic but rather to provide a qualitative numerical description of how the Sharpe ratio varies with information and the overconfidence parameter. Figure A1 plots $Sharpe^K$ as a function of x^K for the increasing degree of overconfidence starting with $K = 1$.

We see that the true Sharpe ratio of the overconfident investor is strictly lower than the Sharpe ratio of the rational investor and decreasing in the level of overconfidence. The sensitivity to the amount of information is also lower the higher the degree of overconfidence. Furthermore, for K sufficiently large, the Sharpe ratio is negatively related to the amount of information at all levels of information. In our computations the relation between the Sharpe ratio and information becomes negative when overconfidence is such that the investor's perceived standard deviation of returns is half its true value.

Figure 1: Sharpe ratio

Notes: The figure plots the relation between investment in information and the portfolio expected Sharpe ratio for the rational investor and for investors with different values of the overconfidence parameter. Calculations are made calibrating the model with the same parameters used by Peress (2004): CRRA utility with relative risk aversion equal to 5, variance of stock returns equal to 2.75%, and equity premium of 6.5%. The relation between the expected Sharpe ration and investment in information becomes negative when overconfidence is such that the investor's perceived standard deviation of stock returns is half its true value.



B: Data sources and variables' definitions

The Unicredit Survey

The Unicredit Survey of Investors' Behavior (UCS) draws on the population of clients of one of the two largest Italian banks. The sample includes 1,834 individuals with a checking account in one of the banks that are part of the Unicredit Group. The sample is representative of the eligible population of customers, excluding customers less than 20 years old or older than 80, and those who hold accounts of less than 1,000 euro or more than 2.5 million euro. UCS goal is to study retail customers' behavior and expectations. The survey has detailed information on households' demographic structure, wealth (both within and outside the bank), and income. It has data on multi-banking, attitudes towards saving and financial investment, propensity to take financial risk, retirement saving and life insurance. Interviews were administered between September 2003 and January 2004 by an Italian leading poll agency, which also serves the Bank of Italy for the Survey on Household Income and Wealth (SHIW). Most interviewers had substantial experience in administering the Bank of Italy SHIW, which is likely to increase the quality of the data. The Computer Assisted Personal Interview (CAPI) methodology was employed for all interviews. Before the interview, each customer was contacted by phone.

The sampling design is similar to that of the Bank of Italy SHIW. The population of account holders is stratified along geographical area of residence (North-East, North-West, Central and Southern Italy), city size (less than 30,000 inhabitants and more), and wealth held with Unicredit (as of December 31, 2003). The questionnaire was designed with the help of field experts and academic researchers. It has eight sections, dealing with household demographic structure, occupation, propensity to save, to invest and to risk, individual and household financial wealth, real estate, entrepreneurial activities, income and expectations, life insurance and retirement income. The wealth questions match those in the Bank of Italy SHIW, and allow interesting comparison between the wealth distributions in the two surveys.

An important feature of the UCS is that sample selection is based on individual clients of Unicredit. The survey, however, contains detailed information also on the household head - defined as the person responsible for the financial matters of the family - and spouse, if present. Financial variables are elicited for both respondents and household.

Construction and definition of wealth

UCS contains detailed information on ownership of real and financial assets, and amount invested. Real assets refer to the household. Financial assets refer to both the account holder and the household. For real assets, UCS reports separate data on primary residence, investment real estate, land, business wealth, and debt (mortgage and other debt). Real asset amounts are elicited without use of bracketing.

Two definitions of financial wealth are available. One refers to the individual account holder, and the other to the entire household. The two can differ because some customers keep financial wealth also in different banks or financial institutions (multi-banking) and/or because different household members have different accounts.

Calculation of financial assets amounts requires some imputation. First of all, respondents report ownership of financial assets grouped in 10 categories. Respondents are then asked to report financial assets amounts; otherwise, they are asked to report amounts in 16 predetermined brackets and if the stated amount is closer to the upper or lower interval within each bracket. The questions are the same used in the Bank of Italy SHIW.

Expected return, standard deviation and Sharpe ratio

To construct the portfolio Sharpe ratio we rely on Pelizzon and Weber (2005), who further classify the 10 UCS asset categories in short-term government bonds (considered to be the risk-free asset), medium-term government bonds (MTGB), long-term government bonds (LTGB), and stocks, as explained in the table below. The questionnaire does not contain exact information on the maturity of government bonds, and the composition of mutual funds and managed investment accounts. Even if the precise split is not known, the survey asks if mutual funds are predominantly stocks or bonds, and we can combine this information with aggregate data to reclassify mutual funds and managed investment accounts.

Asset type	Fraction with positive amount of the asset	Reclassified asset category in the UCS
Bank accounts	94.1	Risk-free
Repurchase agreements	4.9	Risk-free
Certificate of deposits	7.9	MTGB
Government bonds	28.8	Risk-free, MTGB, LTGB
Corporate bonds	27.7	MTGB
Derivatives	2.9	Stocks
Sharesof listed companies	39.4	Stocks
Shares of unlisted companies	3.1	Stocks
Mutual funds	41.4	MTGB,stocks, risk-free
Managed investment accounts	23.3	MTGB,stocks, risk-free

Note. The table reports the reclassification of the assets in the UCS in three asset groups: risk-free, medium term government bonds (MGTB), and long-term government bonds (LTGB).

We estimate the proportion invested in stocks using the average portfolio allocation of Italian managed funds in the 2004 Assogestioni Technical Report. For those who state that mutual funds or managed investment accounts are mostly stocks we assume that 88.61% is invested in stocks, 1.47% in bonds, 9.92% in the risk-free rate asset. For those who state that they are equally distributed between stocks and bonds, we assume that 43.07 percent is invested in stocks, 49.56% in bonds, 7.37% in the risk-free rate asset. For those who state that they are mostly invested in bonds, we assume that 1.55% is invested in stocks, 93.3% in bonds, 5.2% in stocks. Government bonds are allocated according to the composition of Italian public debt: 55% short-term bonds, 1% medium-term bonds, 54% long-term bonds.

Pelizzon and Weber then estimate the first and second moments of asset returns. Holding period returns for short term government bonds are computed from the 6-month Treasury Bill rate, assumed to be the risk free rate. For MTGB the holding period returns is a weighted average of holding period returns of medium term government bonds (80%) and corporate bonds (20%). The holding period return of medium term government bonds is derived from the RENDISTAT index assuming a duration of two years. For corporate bonds we use the RENDIOBB index (the index of Italian corporate bonds yields) and a duration of three years. For long term bonds we use the estimated term structure of interest rates and a duration of five years. All returns are net of withholding tax, on the assumption that for most investors other tax distortions are relatively minor (financial asset income in Italy is currently subject to a 12.5% withholding tax). Stocks returns are computed from the MSCI Italy Stock Index total return.

The sample period is 1989-2003, because some assets did not exist prior to 1989. Pelizzon and Weber exploit the convergence process of Italian interest rates to German rates that accelerated dramatically before the introduction of the Euro in January 1999. Using Weighted Least Squares, the early return series are down-weighted more the farther away they are from November 1998, and weight one after November 1998. The weights are a geometrically declining function of the lag operator multiplied by α, with α equal to 0.8. The weighted series is used to compute sample first and second moments reported below.

	LTGB	MTGB	Stocks
Excess returns %	1.740	0.945	2.179
Standard deviation %	4.271	2.155	20.231
Correlation matrix	1	0.948	-0.194
		1	-0.127
			1

Note. The table reports excess returns, standard deviation, and correlation matrix of medium term government bonds (MGTB), long-term government bonds (LTGB) and stocks. The return on the risk-free asset is 0.9275 percent.