PhD in Economics XXV cycle

Tell me your portfolio and I will guess who you are: social incentives for more fitting pension funds

Alan Giovanni Improta

Supervisor

Prof. Giuseppe Ragusa

Ph.D Coordinator

Prof. Pierpaolo Benigno

Tell me your portfolio and I will guess who you are: social incentives for more fitting pension funds

Alan Giovanni Improta

Abstract

Taking as sample, data obtained directly by the pension fund of an Italian multinational containing more than 35 thousand members, it is assessed, through logistic regression models, how demographic characteristics might affect individual risk aversion. The test is useful to identify groups of workers that by nature are more risk averse and could be disadvantaged by the 2006 TFR (severance indemnity) Italian pension reform. For example women controlling for age, income, region and financial literacy prefer lower risky portfolio and they are more likely to switch toward safer sub-funds.

This analysis could support the policymaker to calibrate a suitable appendix to the last TFR reform in order to cover gaps in opportunities among different kind of risk takers mitigating the so called "social security risk".

In the meantime, it is taken the occasion of such a rich dataset to exploit this sizeable shock in order to test forced (or semi-forced) participation, confirming higher risk aversion for forced participants.

JEL-Classification: J16 (Economics of Gender), G11 (Investment Decisions), H55 (Social Security and Public Pensions), D14 (Household Finance)

Keywords: gender, risk aversion, complementary social security, forced participation

Contents

Section 1. Introduction

Risk preferences are highly heterogeneous. In the whole microeconomics literature, it has been particularly important to analyze the shape of the individual risk preferences in order to define the utility curve. A lot of studies examined factors shaping different risk preferences that are at the base of the investment decisions. Some studies highlight a high impact factor: gender. Many others examine what variables mostly affect the decision to invest in more risky or less risky portfolios.

Different leading papers have conjectured that women are more risk averse than men. But the issue is not so clear as it appears because most of the studies are based on a survey approach or an experimental one. In literature can be identified two different approaches to measure household attitudes toward risk: 1- the revealed preference approach infers risk aversion using weight of risky share contained in portfolios chosen by investors. This is the case of this paper, that takes origin from Olivares et al (2008) where differences among gender are studied throughout a panel data for two different periods and controlling by age, monthly income, balance account and region. 2- The other is based on the elicitation of risk preferences from direct experiments and answers to survey questionnaires.

The aim of this paper is to understand the consequences of a sensible public reform, such as TFR pension reform in Italy and the effect on certain categories of individuals. The Italian reform on TFR ("Trattamento di fine Rapporto") provoked a very important shock in the pension system. In fact, until 2007, a worker received a gross wage, from it the employer held a part to pay taxes and his/her consumes, a part is taken as contribution for the employee's pension and a part as contribution for an eventually lump sum that the employee will collect from the firm at the end of the working relationship. This last contribution is called TFR (Trattamento di Fine Rapporto) and was collected in the firm balance sheet until the employee left the firm. It is not a redundancy pay-out, since he/she will get it even if he/she resigns voluntarily. It is like an "end-of-service pay-out". The amount of it is about the total wage divided by 13.5. So, it is a huge part of the personal income. It represented a low cost resource of financing. The firm had to compound it at very low rate; according to the Italian Law at 1.5% + (0.75 * inflation rate). At the end of

2005, the Italian Government formulated a very revolution in the field of the pension system. It established that starting from 1st January 2007 the employee had to choose if holding his TFR in the company or to bring it to a chosen pension fund. For companies with more than 50 employees, even if the employee decides to leave his/her TFR in the firm, the firm is obliged to transfer the contribution to a pension fund. Due to the reform also agents with scarce information about finance had to choose about the category of risk fund, a clear example of forced participation. According to previous studies, females are more risk adverse. We want to effectively prove whether effectively Italian women are more risk adverse than men when choosing different retirement plans based on risk preferences. The social issue is linked to the fact that since women choose lower risk funds (that in the long run should yield less, as in figure 2), earn lower wages (and so make smaller monthly TFR contributions) and work for less years, they will accumulate a lower end-of-service pay-out, that moreover it should be used for more years (since average life for women is higher than for men). The consequence could be an insufficient flow to assure an adequate living standard (also called social security risk).

To analyze the Italian case, I take as sample data obtained directly from the Fopen, the pension fund of the Enel Group (and other ex-Enel companies) with 37366 subjects under analysis. This fund gives the opportunity to each member to choose among 6 categories of portfolios (only one till 2003 and only four till 2007). The difference among types of funds is the investment limit in stocks. The investment limit in stocks in each fund is: 70% for the riskiest (Prevalentemente Azionario), 50% for the second risky fund (Bilanciato), 30% for the third (Bilanciato Obbligazionario), 10% for the fourth (Obbligazionario), 0% for the lowest (Monetario e Monetario Classe Garanzia). To my understanding, no study has been undertaken on Italian pension funds focusing on gender variable and fund choice. Moreover the dataset is useful not only to extract information about factors affecting aversion to risk but also to model the dynamic of the choice. A natural shock helped us also to understand the choice for individuals forced to participate. Anyway in the period of the analysis (from 2000 to 2012), 2 interesting points of time are been used as shocks: 2003 since the pension structure went from 1 risk category to 4 categories, then 2007 when an invasive reform has been implemented.

In the first kind of test, using a Binomial Logit model we find the main factors affecting the choice to join the most risky category or the less risky one. Since data are standardized, it is easy to interpret the results as score of contribution to opt for high risk (less risk aversion) or low risk (more risk aversion) categories. Results clearly highlights that gender, age, income, geographical factor and trading really affects aversion to risk in line with the previous literature like in Bajtelsmit and Vanderheid (1997), Guiso and Paiella (2009) etc. That is, female are more risk adverse also controlling for many variables. Huge evidence of factors provoking more risk aversion are advanced age and living in less developed geographical areas (like South or Islands). While propensity to trading, higher income and living in more developed areas are factors affecting risk aversion negatively. Wealth is the only ambiguous factor, but it could be due to weak construction of that variable (since the observation of such a variable is only partial). Moreover, using the same set of data and adding a new dummy variable capturing the date of inscription of each individual before or later an established date, it can be inferred that people with lower interest in financial matter (or less financial knowledge, or merely forced to participate) are more risk adverse even controlling for other factors.

In the second kind of test we go more deeply to the dynamic of the choice, that is, the aim is exploring factors that make the switch toward an higher risk category more or less likely. Using as in the previous test a Binomial Logit model I take into consideration a very large set of switches during the period from 2003 to 2012. This kind of test confirms the first one: women are more likely to change toward a lower risk category even controlling for age, income, trading, wealth and geographical factors. It appears even clearer if we carry on the analysis only on the sample of employees that had to switch in 2003 from the only one available category to one of the new 4 categories.

In a third test, throughout test of hypothesis, as in Olivares et al (2008), we verify the existence of differences in the weight of men and women that change their portfolios within fund categories. This allows to eliminate the notion that part of previous results of the paper are driven by the number of women and men in the sample. When controlled by proxy of income and age, the difference not only persist but slightly increases. In addition, with the same method, it is demonstrated that forced participants (or anyway workers that joined the fund before the reform) are more risk adverse.

The drawback of the previous literature (using experiments or choices under shocks) ridden over by this paper is that: here behaviors are not elicited in hypothetical setting but reflect individual risk attitudes in actual financial decision. While the main drawback in the revealed preference approach, that this paper is able to overcome is about the endogeneity of the variable wealth (and income). That is, in previous papers, if an household benefits a pay rise, unless the household rebalances its portfolio immediately, its financial wealth increases and its portfolio risky share mechanically shrinks. But in the case of this paper, individual chooses a fix percentage on his/her income, that implies a constant share of risky assets unless he/she changes to another risk category. Under this prospective we do not exclude the presence of DRRA[1] (decreasing relative risk aversion). But anyway, the aim of the paper is to encourage the lawmaker to consider a useful appendix to the previous reform, giving the opportunity to disadvantaged agents (by nature) to take decision starting from the same risk attitude. Otherwise, the reform stand alone is giving the opportunity to freely take advantage of the market only to a determined class of agent (the less risk adverse).

The paper is organized as follows. In Section 2, I will introduce the literature review on this topic, where, introducing the main results, drawbacks about different kind of studies are examined. In Section 3, the Italian pension system will be described in order to introduce the important reform of 2007 taken as a shock; moreover, a summary of Fopen history and sample data will be introduced. In Section 4, methodology will be discussed in order to show the numerous results. Finally in Section 5, I will comment the results in the light of the TFR reform.

---

[1] Risk preferences are highly heterogeneous: $\gamma_i = \frac{\lambda_i}{\omega_i^{\eta}}$ Where λ_i is an individual fixed effect that captures unobserved risk preferences, ω_i^{η} is the effect of wealth due to η. η can assume different values: $\eta = -1$ in presence of Constant Absolute Risk Aversion preferences (CARA); $\eta = 0$ means Constant Relative Risk Aversion (CRRA); $-1 < \eta > 0$ in case of Increasing Relative Risk Aversion and Decreasing AbsoluteRisk Adversion; $\eta > 0$ in presence of Decreasing Relative Risk Adversion (DRRA) and Absolute Risk Aversion.

Section 2. Literature Review

The literature in this field followed mainly 2 approaches: the first is based on a revealed preference[2] strategy that infers risk aversion from the portfolio risky share chosen by investors in real life. The second relies on the elicitation of risk preferences from subject behaviors in experiments and answers to survey questionnaires (Guiso and Sodini, 2012).

This study is focused on highlighting that gender plays a role in the individual/household portfolio allocation decision, where women tend to invest in portfolios characterized by a lower level of risk (volatility) and clearly they are more risk adverse then men. Literature on gender difference in risk aversion is further divided into 2 lines. The first focuses on finding if there is actually a gender difference in risk aversion. The second one focuses on the psychological factors that would result in women being more risk-adverse than men. In my case, what is appealing in the study is to investigate if women are more risk-adverse than men after controlling for wealth, income, age, trading and regional factors. Moreover, it will be interesting to know if other categories of subjects are more sensitive to risk aversion. This is important for social implication as Sunden and Surette (1998) suggest: gender differences in investment decisions exist, but they are more complicated than previous literature have suggested, including also marital status. They moreover infer that if women are not making optimal investment choices this could severely impact their accumulated wealth for retirement.

As written before, methods to extrapolate and analyze risk aversion in individual investment behavior space from the use of data on actual assets holdings to questionnaires soliciting hypothetical portfolio decision. Results varied according to the dataset used and due to the control variables used. This is the first paper in the household finance literature using large Italian actual data describing the concrete choice of individuals in defined contribution plan (also if only for the part

---

[2] Revealed preferences approach is based on the intuition that $\gamma_i = \frac{Er_i^e}{\omega_i \sigma_i^2}$ where Er_i^e is the expected risk premium, σ_i^2 is the return volatility of risky assets and ω_i the fraction of financial wealth invested in risky assets (portfolio risky share)

regarding the so-called TFR contribution). Many studies have largely used the Federal Reserve's Survey of consumer Finances[3] (SCF), which is a triennial survey of the balance sheet, pension, income, and other demographic characteristics of U.S families. Sunden and Surette (1998) use data from this survey, as well Jianakoplos and Bernasek (1998) that find that single women exhibit relatively more risk aversion in financial decision making than single men. They examine household holdings of risky assets to determine whether there are gender differences in financial risk taking. As wealth increases, the proportion of wealth held as risky assets is estimated to increase by a smaller amount for single women than for single men. Gender differences in financial risk taking are also influenced by age, race, and number of children. Greater financial risk aversion may provide an explanation for women's lower levels of wealth compared with men's. Also Bajtelsmit, Bernasek and Jianakoplos (1999) use this kind of data, noting that in view of the longer life expectancy of women, even given the same investment strategy and retirement savings as men, consumption in retirement will be less for women. Thus if greater risk aversion is evident in retirement saving decisions, women's consumption levels could be eroded.

Other datasets (but regarding Italian individuals), are the UCS survey that is conducted on a sample of Italian individual investors owning a checking account at Unicredit, a large European banking group and the Italian Survey of Households Income and Wealth by Bank of Italy). Respectively used by Guiso, Sapienza and Zingales (2011) and Guiso and Paiella (2008). The first paper contributed to suggest that risk aversion does fluctuate in a major way. Hence, it is possible that fluctuations in risk aversion can explain those movements in asset prices that are not justified by changes in expected cash flow. In fact, they document that individual risk aversion increases substantially following the 2008 financial crisis. This increase cannot be explained on the basis of standard reasons (such as changes in wealth, habits, or background risk). The only variables that have any explanatory power are proxies for changes in confidence. Moreover, they test that fear can have an heavy impact on changing the risk aversion, but they do not test for gender differences. While the second paper tests in the reality the hypothesis of CARA preferences through a direct measure of absolute risk aversion based on the

[3] The pioneers in using this data in the risk aversion field are Friend and Blume (1975)

maximum price a consumer is willing to pay for a risky security. They find that consumer's environment (background risk) affects risk aversion more than taste and demographic parameters.

Sometimes the size of the sample analyzed has been dramatically small, like in Bernasek and Shwiff (2001), where the sample included only 270 faculty. They find that gender is a significant factor explaining the difference in the percentage of an individual's retirement fund invested in stocks, with results conforming to the theory, that is, percentage decreases as respondent are women.

As in Bernasek and Shwiff (2001), interesting studies focused on the importance of cultural background as factor to control for gender difference. This kind of studies try to verify the so call hypothesis of "expertise dominates gender". There is some evidence that decisions in financial affairs may differ from decisions in abstract gambles, possibly because financial decisions involve clear incentives. In detail there seem to be two separate forces which reduce the gender difference in risk aversion, i.e. familiarity with risk and risk decision under financial framing. A clear example of this studies is by Beckmann and Menkhoff (2008) that conducted a written survey with professional fund managers in the United States, Germany, Italy and Thailand between spring 2003 and winter 2004. Testing the "expertise dominates gender" hypothesis surprisingly ends in a victory for the gender difference. In fact, controlling for a large set of competing influences, the gender variables always shows the sign as expected from the earlier literature, that is female fund managers keep their more risk adverse behavior but on the other hand the effect is comparatively weak for the established risk measure. Moreover, they find that the relative economic importance of the gender-related difference in explaining behavior is sometimes small in comparison to competing influences, indicating that indeed financial expertise decreases the gender difference, but does not erase it.

Instead, in Hibbert, Lawrence and Prakash (2008), measuring gender difference in risk aversion using a sample that controls for biases in the level of education and financial knowledge, they conclude that when individuals have the same level of education irrespective of their knowledge of finance, women are no more risk averse than men. They used a dataset resulting of a survey of Finance and English faculty at universities in the US. Since all individuals in their sample have achieved at least a graduate degree, they implicitly control for the level of education. Finally, they conclude that gender difference in risk aversion is confined to individuals who are married or live with a partner, in fact, single women are no more risk-adverse than single men when they are both highly educated. But, analyzing the class of assets considered the riskiest, their results are in line with those of my paper, that is, when both men and women invest in the asset class they consider most risky, women are more likely to invest the smallest portion of their portfolio in that asset class.

In the case of my paper, even if it is possible to infer the role of cultural level on the aversion to risk, it does not matter because the aim of the paper is to identify more sensible categories that cannot enjoy the whole benefits of the TFR reform, independently of their financial expertise. It would have been interesting to explore the size of the role of financial knowledge only if the possible relief of giving everybody an elevated culture (or a sufficient financial knowledge) would be feasible, but it is clearly too expensive both by a social and an economic view.

In the case of studies done by surveys the main drawbacks are the following. First, when asked about willingness to pay, individuals tend to underreport, which overestimates their true risk aversion (Kachelmeir and Shehata, 1992). Second, answers may be affected by how questions are framed. Third, the validity of this methodology rests on the assumption that respondents know how they would behave in a hypothetical settings and that they are willing to reveal truthfully their choices (Kahneman and Tversky, 1979).

Also experiments has been used frequently, but in most of the cases the size of the sample is even smaller. For example Felton, Gibson and Sanbonmatsu (2003) examine the role of gender and optimism in determining the attitude for risk in investment choices of 66 undergraduates students with both monetary and academic incentives.

Grable, Lytton and O'Neill (2004) survey 421 relatively young individuals (average age of 32.03) via the internet and find that men report a higher risk tolerance score than women. Even though gender difference was not the focus of their study (that investigates if projection bias, as explained by regret theory, shape financial risk tolerance attitudes), they find that gender plays an important role in explaining the attitude toward taking investment risks. However, their study does not control for other variables (considered in my paper), such as income and age which are known to explain risk-tolerance.

Compared to surveys, it is however more difficult to link lab experiment findings to actual behavior outside of the lab, partly because subjects are typically students who typically have not yet faced actual financial decisions, partly because they often are selected samples not representative of the population (Guiso and Sodini 2012). This paper can be considered as a kind of large experiment, but results are not elicited in hypothetical setting but in real life choices. So it reflects individual risk attitudes in actual financial decision.

If we look at relevant studies or comments of the complementary pension funds issues, it is appropriate to mention a study by the Bank of Italy that expresses worries about the size and the diffusion of the risk that the employees' cumulated saving is likely to result insufficient to finance the consumption in the retirement period. Beyond the mentioned TFR reform, few years before the Italian pension system changed radically, moving from total wages system ("sistema retributivo" to contributive system also called "pay-as-you-go"). This change implied a remarkable fall of the pension size. The social security risk (that we call also pension risk) is higher for those workers that not only will suffer a fall in the main pension, but also for those having a wealth and saving not suitable.

The valuation of the suitability of the saving requires exercises that go beyond this discussion (that is focused on TFR reform), and should take into consideration both the overall wealth of each individual at time of retirement and the needs of consumption in the following years. To go deep in this, it is useful to read Skinner (2007) and Fornero at al. (2009).

In any case, the study published by the Bank of Italy (Cappelletti and Guazzarotti 2010) shows that the social security risk is mainly diffused among youths, employers from South and Island of Italy, employees of the private sector and workers in the lowest income clusters. The same study evaluates an empirical model that analyses the choice to join complementary pension funds according to many socio-demographic variables, substitution ratio between pension and last income, and dummies about financial literacy and pension literacy. The most significant explanatory variables are the income and level of instruction of the breadwinner. While the age of the worker encumbers on the probability to join complementary pension fund negatively.

To have an idea of the size of the phenomena, the Minister of the Economy of Italy estimates that in case of a private employee depositing the complete TFR in a pension fund, 30 years later (according to a yield after tax of 3%) he/she will receive a complementary pension of about 15% of his/her last income (that is 1/5 of the overall pension income). Starting from this, workers choosing for 30 years a high risk fund could receive till 25% of his/her last income, making of really important interest the choice of the category of risk.

According the Bank of Italy, from 2002 to 2008, 70% of workers chose a low category of risk, while the remaining part opted for a high risk category. In any case the size of the sample used for that study is not enough to built a statistically significant model in order to analyze the risk aversion. But their simple statistics show higher propensity to the risk for younger workers.

In 2010, another study (Cappelletti, Guazzarotti and Tommasino, 2010), belonging to the paper series of the Bank of Italy, takes into consideration a small Italian sample that differs largely from the Italian population. In fact it differs largely from the survey on household income and wealth (SHIW), commonly used in Italian papers on Household Finance. In the study they estimate a multivariate probit model that takes into consideration the fund level of risk joined by about 3800 people mainly clerical or managerial workers. They focus on the effect of age on the risk aversion, after controlling for some socio-demographic characteristics, and find a pronounced tendency to choose safer portfolios as people age. Moreover, they try also to model characteristics that make participation in pension funds more active, analyzing the switches. The paper, at first sight could be defined similar to what I provide. But apart very different datasets and econometric models used, the aim of the two studies is extremely different. They suggest that life cycle funds could be a valuable instrument, given that they automatically bring all the participants toward less risky allocations as they get near to retirement. While I focus on characteristics making workers by nature more risk adverse. Furthermore, I use the switches and the flag trading in order to verify conditions making the individual more or less risk adverse. In few words, I focus on the opportunities lost by such people due to the last TFR reform. I explain that, utility maximization at time t in choosing a certain category of risk is not the same than choosing a category at time t+n. In fact, most of the workers will not enjoy enough pension to provide their consume. This can strengthen the so called social security risk, that can lead to higher public disbursement to safeguard retirees wealth.

Section 3.1 The 2006 TFR reform in Italy and brief history of Fopen

This study took inspiration by the pension reform that in 2007 shocked millions of Italian workers. It really changed the way the severance indemnity (TFR, Trattamento di fine rapporto) was accumulated. TFR, or severance indemnity is a payment due to workers upon termination of their employment. In Italy, this measure was introduced back in the early 20th century for social security but also for welfare reasons on account of the loss of the employment. Originally, it was paid to white collars only, and progressively it was also extended to blue collars. Severance indemnity is now regulated under article 2120 of the Italian Civil Code. It is calculated on an annual basis and is equal to 6.91% of the remuneration. Upon termination of employment, the indemnity amount is revaluated with a rate equal to 75% of the inflation rate, plus 1.5% as fixed rate. At fiscal level, a distinct taxation regime applies to severance indemnity (previously not taxed): the rate is equal to the average rate for personal income tax over the previous five years (with a minimum threshold of 23%), while revaluations are taxed at 11%. From January 1st 2007, companies with over 50 employees must pay the severance indemnity, not transferred by the worker to supplementary pension schemes, to INPS (the national pension scheme). However, this does not affect the workers' rights, as they will continue to be entitled to, as prescribed by law (termination of employment, early termination, etc.), a severance indemnity from their companies, which will then recover such amounts from INPS.

The reform pushed workers toward a totally new culture: from a world in which pensions were guaranteed by society as a whole to one where they are ever more resting on one individual's informed participation in choosing how much to invest and which balance to strike in the trade-off between risk and return. Due to the

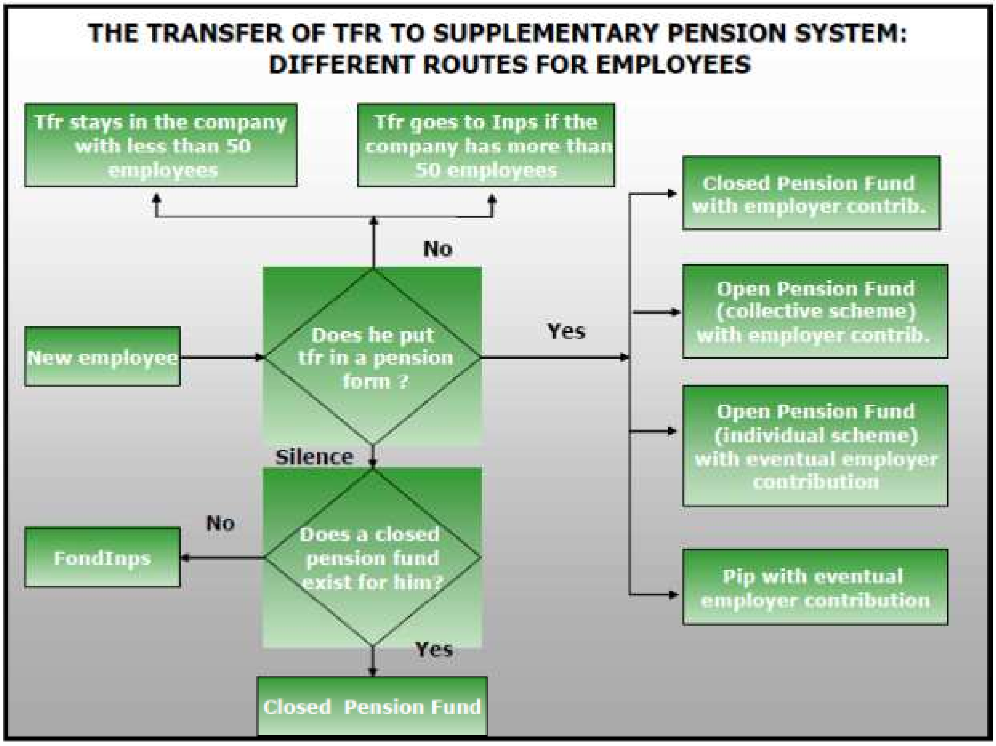
2005 reform, within 30th June 2007, or within 6 months from recruitment, when dated after the 1st of January 2007, workers may:

- keep their severance indemnity in their company;

- transfer it to specific supplementary pension schemes;

- not take any decision.

Figure 1



In the first case, companies with over 50 employees must pay the severance indemnity, not transferred by the worker to supplementary pension schemes, to INPS (the Italian Public Social Security Institute).

In the second case the employee has to choose a specific complementary pension form where to put TFR: collective (Closed/Open Pension Funds) or individual (Open Pension Funds or PIPs). In the collective schemes contributions come from both employer and employee, as established by agreements. In the individual schemes employer's contributions are optional.

In the former case, the severance indemnity, as it accrues, will be transferred by the employer to the pension scheme indicated in the collective labour contracts or, failing that, in a specific supplementary pension scheme called Fondinps set up by INPS. The decision to transfer the severance indemnity to a pension scheme may not be revoked, whereas the decision to keep the severance indemnity in the company may be changed with a view to join a supplementary pension scheme.

In my analysis I used this big shock in order to test risk aversion among individuals that liked to transfer their TFR to Fopen. In fact, data has been taken by the collective closed pension funds of the Enel Group, a fund chosen by more than 40 thousand people in the last years, that is one of the biggest in Italy, managing more than 1 trillion euros.

The history of Fopen is a little longer than the TFR reform. In fact it was born in 1999. Anyway, the first collection began on the 16th November 2000, but till the 16th April 2002 the fund invested only in short-term bonds (free of risk investment). Since then, there has been the introduction of a unique risky category (called "Unico"), where 62% of the total was invested in bonds and 38% in stocks, a rather risky portfolio. Only since the 16th of June 2003 has been introduced the opportunity of choice among multiple categories (4 different level of risk: "Monetario", "Bilanciato Obbligazionario", "Bilanciato" and "Prevalentemente Azionario". On the 1st of April 2006 a fifth category has been added ("Obbligazionario") and finally in August 2007 a Monetary similar category has been added ("Monetario Classe Garanzia").

To sum up, at the time of data (May 2012) workers can choose among 6 different risk categories (but with only 5 levels of risk). The investment limit in stocks in each fund is : 70% for the riskiest (Prevalentemente Azionario), 50% for the second risky fund (Bilanciato), 30% for the third (Bilanciato Obbligazionario), 10% for the fourth (Obbligazionario), 0% for the lowest (Monetario e Monetario Classe Garanzia).

The choice of joining the FOPEN is subsidized by Enel through a money incentive in the size of the 1.35% of the gross wage. So that the majority of employees prefer to join this fund instead of the public one mentioned before. This is important because allow us to examine even subjects slightly more risk adverse. Moreover, many workers joined the Fopen voluntary, much earlier than the reform passed or has been even discussed. In the next sections this difference will be exploited in order to have a kind of control group (we can assume that who enrolled to the Fopen in earlier stage are more interested in finance or at least they have more familiarity with financial decisions).

Section 3.2 The Dataset

The source of this paper is composed of 3 main dataset (see figures 3,4 and 5): “soci”, “switch” and “quote socio”. The first set contains more than 43 thousand anonymous but coded individual. It contains the following information: individual’s code, year of birth, gender, kind of participation, date of enrollment, province of residence, region, percentage of voluntary personal contribution, contribution by the firm, percentage of tfr versato in the the fund, personal contribution in euro and contribution of tfr in euro.

The second set called “switch” contains more than 59 thousand changes of categories, done by anonymous but coded individuals. It is composed of these information: individual’s code (that can be linked to the first dataset), year of birth, date of enrollment, current status (if active or left), date of change of risk profile, cumulated contribution at time of change, name of category of origin, name of category of entry, number of shares, value of a share.

The last set, “quote socio”, contains the following columns: individual’s code (that can be linked to the other datasets), current category (updated to May 2012), number of shares, value of a share, cumulated contribution.

Most of the variables are transformed in categorical variable, and then transformed in many dummy variables. The variable “age” is divided into 4 categories: A- under 36 years old, B- between 36 and 46 years old, C-from 46 to 53 years old and D- over 53 years old. The variable “region” is composed of 5 types: A- islands, B- south, C-Centre, D- north-east, E- North-west. The variable “trading” is a dummy that identifies with 1 who change category more than once. The variable “income” is obtained dividing the voluntary contribution in euro by the voluntary percentage of income and then divided into 4 categories: A- up to 1900€, B- from 1900 to 2500€, C- from 2500€ to 2900€ and D- over 2900€.

Looking at the figures 6 and 7, it is easy to extrapolate some descriptive statistics about our sample of employees (37366 still active subjects, even if Fopen has accounted for more than 50 thousand individuals since 1999). Males are the preponderant part of the dataset[4] with about 83.6% (31227 men) of the sample. The distribution of the sample among Italian region is more equilibrated (Enel has been the monopolist in the Energy sector till in the 90s). In fact 20.6% of employees live in the South, 13.3% in the Islands, 27.5% in the Centre, 15.6 in the North-East and the 23.1% in the North-West. With respect the usual Italian population distribution, it appears a little predominance of the Centre Region due to the presence of the Enel Headquarter.

Since the first sight to the data (see figure 8), it is evident that the percentage of subjects within regional clusters choosing for riskier categories such as "Prevalentemente Azionario" or "Bilancianto" is higher in North-East (36.54%) and Centre (31.85%) areas. Moreover within gender clusters is easy to notice differences in preference between men and women (see figure 9): while 13.85% of women choose the less risky categories ("Classe Garanzia", "Monetario", "Obbligazionario"), only the 10.5% of men do the same choice, with a gap of 3.35%. In the same way, if we look at the part of women choosing for the highest risk category ("Prevalentemente Azionario") the gap between women and men is 5.69% in favor of men. That is, the percentage of women choosing for this class is only 7.49%, against the 13.18% of men. Looking at other general statistics in figure 10, we see which percentage of subjects within their age group opts for riskier funds. It is evident the downward trend as the age group goes up. For example the percentage of workers in the cluster "under 36" the choices for the 2 highest categories is 44.85%, it slows down to 36.43% for the cluster "36-46", 28.51% for the group "46-53" and finally reaches 18.89% for the oldest group "over 53". Analyzing the sample under the "income" view, it is marked the preference for more austere income towards less risky categories. In fact, as seen in figure 11,

[4] In my tests I finally consider (in session 4.4) also this unbalance towards men using the same technique of Olivares and Al. (2006) taking as point of measures the change in proportion within gender variable.

16.17% of workers earning less than 1900€ per month prefer less risky categories such as ("Classe Garanzia", "Monetario", "Obbligazionario"), while among who earn more than 2900€ per month the percentage slows down to 9.7% (exhibiting a monotonic decrease in this case as well).

It is very important to underline the fact that the younger is the worker the less she/he is likely to have an high income (see figure 12 and 13). If we see the composition of the group of employees earning up to 1900€ per month we see that the vast majority is made of workers younger than 36 years old. While if we see the composition of the group of workers older than 53, they earn mostly between 2500€ and 2900€ or even more than 2900€. This detail is to be taken into consideration under the light of the following econometric study (section 4.2). In fact, it means that to be young lead toward a more risky category, but to be young usually implies to earn less as well and so, choosing for a less risky fund. The 2 effects behave in opposite direction, making it more interesting to discover the power of them. Anyway this help the model to avoid trivial conclusion and avoid heterogeneity issues (the 2 factors in this way are well decomposable).

Section 4.1 Econometric model

In the first two analysis of this paper, after a careful consideration, I decided, as in Haliassos and Bertaut (1995), to use a logistic regression model in order to discover the main variables affecting the size of risk aversion. A logistic regression, also called a logit model, is used to model dichotomous outcome variables.

Workers i choose to join a determined category of risk sub-fund if unobserved relative indirect utility $V' \equiv V_{SubfundA} - V_{Othersubfunds} > 0$, where

$$V'_i = x'_i\beta + u_i$$

Here, x'_i is a vector of observables, and u_i represents unobserved factors influencing relative utility with a logistic cumulative distribution.

In the logit model the log odds of the outcome is modeled as a linear combination of the predictor variables. Since I tested the risk aversion through the preference of each individual (more than 37 thousand observations) in choosing a risky category or not risky category, or the willingness to switch for a less or more risky category, the dependent variable is structured as binomial.

It is not the only one way to describe conditional probabilities. Another common method is the classic OLS regression. This model, when used with a binary response variable, it is known as a linear probability model. However, as first drawback, the errors from the linear probability model violate the homoscedasticity[5] and normality (because the criterion has only two values) of errors assumptions of OLS regression, resulting in invalid standard errors and hypothesis tests. The second disadvantage is that the OLS model does not properly restrict the range of the dependent variable to the unit interval, as it should, because it is supposed to be a probability. Therefore, nonsensical predictions outside the (0,1) interval are possible if extreme values of regressors are considered.

But as we will see later on, an extraordinary advantage belongs to OLS model: parameters can be interpreted directly as marginal effects, and the approximation is good as long as we do not move too far away from the means of the explanatory variables.

An early treatment of the logit model can be found in Berkson (1944) who considered this model in the context of estimating the effect of a continuous treatment on a binary outcome by the subject. This model specifies:

$$\pi_i = G(x_i'\beta) = \frac{exp(x'\beta)}{1 + exp(x'\beta)} \qquad \text{(eq. 1)}$$

And clearly ensures that $0 < \pi_i < 1$. Moreover, in contrast to a similar (and also very use) model called probit, choice probabilities are available in closed-form and it does not need to integrate in order to obtain the probability. Logistic distribution function is flatter than the distribution function of the standard normal. In addition, it is very similar to the probit function. In fact, results of this paper can be easily rescaled in probit form: empirically the factor 0.625 can be used to approximate the parameter estimates in the probit model from the logit one.

[5] y_i can take only two values, zero and one. Hence the implicit regression error $u_i = y_i - x_i^{'}\beta$ can take only two values, $0 - x_i^{'}\beta$ and $1 - x_i^{'}\beta$ so it follows that $VAR(u_i|x_i) = x_i^{'}\beta(1 - x_i^{'}\beta)$, thus the error term is heteroscedastic.

Another useful and very common expedient (in the statistics literature) that the specific structure of the logit model offers, is the odds interpretation.

From eq. 1:

$$\frac{\pi}{1-\pi} = exp(x'\beta) \;\; \Rightarrow \;\; ln\left(\frac{\pi}{1-\pi}\right) = x'B$$

Here $\pi/1-\pi$ measures the probability that Y=1 relative to the probability that Y=0 and is called the odds ratio or relative risk. For example, consider in this case Y=1 denotes the choice of the highest risk category (within the fund) and y=0 denotes any other category and regressors include a measure of income. An odd ratio of 2 means that the odds of choosing the highest category are twice those of picking any other. For the logit model the log-odds ratio is linear in the regressors.

From the literature, I learnt the following consideration, that is, although OLS estimation with heteroskedastic standard errors can be a useful exploratory data analysis tool, it is best to use the logit model for final data analysis.

Section 4.2 Logistic Regressions with standardized data

In the first part of this study I adopt a logistic regression using data contained in the datasets called "Soci" and "Quote Soci". The observed outcome is a dummy linked in one case to belong to the highest risk category ("Prevalentemente azionario"), while in the other case to belong to the lowest risk category ("Monetario"). That is, Y=1 if the employee belongs to the risk section of interest (in one case the highest, while in the other the lowest), Y=0 if the employee does not belong to the risk section of interest (as before). The joint analysis of both cases places the model between the standard binary logit model and the ordered logit model.

After defining the available variables included in the dataset, factors (X_i) are selected through a Stepwise method. At this point it is fundamental to assess the fitness of the model through the percentage of concordant. That is, taken a couple of observation with Y_i=1 and Y_h=0, it is called concordant if $P_i > P_h$, tied if $P_i = P_h$ and discordant if $P_i < P_h$. In the assessment of the fitness I take into account 4 indexes.

Somers'D = $\frac{C-D}{N}$

Gamma = $\frac{C-D}{C+D}$

Tau-a = $\frac{C-D}{(0.5O(O-1))}$

c = $(C + 0.5T)/N$

As those indexes obtain higher score, the higher is the fitness of the model. In fact, higher values correspond to a stronger association between predicted values and observed ones.

Then we look at the joint significance of the coefficients through the Likelihood ratio test, Score test and Wald test. After the joint significance, I checked for the significance of the single coefficients.

Then, the presence of multicollinearity has been checked, analyzing the correlation matrix of the regressors included in the model. In the case of 2 or more regressors strongly correlated, I deleted those presenting a lower correlation with our explained variable[6]. At the end of this selection process I run again the model including only the survival independent variables. After this procedure, the logistic model on standardized data is able to supply scores of sensitivity of the selected factor. At the base of this affirmation there is the standardization process, that allows an objective interpretation of the coefficients of the different factors. They can be compared because they are in the same unit and can be represented on a graph with a common scale. It means, that even if the results belong to 2 different regressions (with a totally different dependent variable), they can be compared and taken as part of the same regression. It is useful not only to look at the coefficients but also to their respective odd-ratios. For the moment it is not relevant the analysis of the marginal effect because of the standardization process. But, anyway,

[6] I look the coefficients of linear correlation between the dependent variable and the whole potential regressors

it is important to highlight the fact that the logit model is not a linear model since $G(.)$ is a non-linear function. Therefore, the parameter $\beta_l$associated with the l-th element in x_i does not directly measure the marginal effect. But we will explain it in detail during the presentation of the second analysis.

Now, it is better to focus on the result of the first regression (called R1a), where $Y = 1$ if employee opted for the highest risk category, while $Y = 0$ for any other. The best model (see table 1) obtained after a strictly selection process includes 12 factors plus the intercept. Analyzing the sign of the coefficients we are able to indicate the following (figure 14 offers a clear view of the results).

Our main variable of interest is gender. Gender dummy scores 0.2921 and it is significant since its p-value is lower than 0.0001. The odd-ratio associated[7] at this score is 1.339 with a limit of confidence between 1.288 and 1.393. This variable behaves in line with most of the literature (and is compatible with what we argue in describing the social issue of the reform), that is, females are more risk adverse than males. This dummy is verified thanks to the introduction of many other control variables.

Since we have the youngest age clusters with a positive sign (0.3139 and 0.1838) and the oldest cluster with a negative one (-0.345), it means that age play a relevant role in choosing a determined risk category. These coefficients are significant, in fact their p-value are lower than 0.0001. It can be easily affirmed that the older is the employee the less likely he/she will choose a risky category. That is, elders are more risk adverse (and scores are monotonic in age).

Also the variable income is present in this regression. In fact, the dummy of the cluster of the richest employee is significant but with a p-value of 0.0092 and its coefficient scores 0.0612 (positive), meaning that richer employees are less risk adverse.

Then 3 dummies refer to the macro-regional areas: positive and significant coefficient for the flag North-east (0.1567), negative and significant coefficients for the flags Islands and South (-0.0736 and -0.1246). These variables can be

[7] At the moment we do not consider the real marginal effect as these regressions are in standardized data

interpreted in the same direction but with 3 different causes. In fact, it seems that living in most developed region of Italy makes subjects less risk adverse. It can be due to economic factors, social factors or cultural factors. Better economic conditions can have an impact on the background risk. In fact, for the Enel employees there is no background risk, but each worker lives in a different economic context and it is more likely that in the less developed regions the partner is unemployed, making the weight of the income coming from the employee more important to the family wealth. Unfortunately, we have no data about the marital status of the subject nor about other kind of family income. But also social factors can have an impact on the sensitivity to take risks. People living in the South of Italy or in the Islands, traditionally grow with stronger bias against risk. Another cause is the propensity (cultural factors) of people living in more developed regions to take studies in economic and financial studies (or anyway a simple higher index of schooling). Whatever is the causes, the model seems to work in the right direction even for these variables.

In this model only a variable does not work as it could: the position accrued (also called accrued balance). We can consider it as a proxy for wealth, or at least it is a form of raw asset property. This factor does not show a coherent behavior because in both regressions R1a and R2a have the same sign. In fact the variable at the same time (it is negative) pushes the worker to not choose nor the highest nor the lowest risk category. This means that it does not work, and as matter of fact it cannot reflect the wealth of the agent, or it means that people with huge accrued balance prefer a category of risk that is in the middle, that is, he/she decides to not choose.

Another important factor, that we will find also later on, is the dummy about trading. This is a peculiar (and original) explanatory variable. An element of innovation in this kind of studies. This variable can be taken as a proxy very linked to the financial expertise. In this regression it works as expected: the coefficient is 0.6202 that is not only positive but also significantly very high. The associated odd-ratio is 1.859, that could impact very much on selecting higher risk categories. This imply that financial knowledge lead people to be less risk adverse. The inclusion in the model of this control variable makes the gender dummy even more important, because it let us to reject the "expertise dominates gender" hypothesis as in Beckmann and Menkohoff (2008).

In the second part of the test I run a second regression (R2a) on the same data. But this time I set Y=1 for subjects belonging to the lowest risk category and Y=0 for the control group choosing for any other category. In this case the best model includes 9 variables (see table 2).

There are many confirmations with respect the first regression (figure 14 offers a clear view of the results). The most relevant is about gender. In fact, in this case the coefficient associated to the dummy is negative (-0.0882) with a p-value lower than 0.0001. That is, men are not likely to choose the lowest category in risk, confirming the higher risk aversion of women. Considering the score, in this case the effect is less marked but compatible with our expectations.

Among the control variables, age works very well. Here we find opposite scores with respect the first regression. This confirms and strengthens our previous statements that elders are more risk adverse after controlling for income. In fact, the coefficient for the youngest cluster (till 36 years old and between 36 and 46 years old) are both negative (-0.2536 and -0.0832), while that for the oldest one is positive (0.1555). Scores show a monotonic behavior in age also in the second regression. So, elders are more likely to prefer the less risky category ("Monetario").

In the second regression there is not income variable. But it is not important, because it would only have confirmed what accepted in the first regression. Here, the dummies referring to macro-geographical areas still prove that living in developed areas makes subject less risk adverse. In fact, in this regression there are 2 significant geographical variables. The flag South has a positive coefficient (0.0821), while the flag North-East has a negative one (-0.0575). The respective p-values are 0.0002 and 0.02. This means that the symmetrical result is confirmed: in the first regression the flag South is -0.1246, implying less likelihood for a South resident to choose an high risk category, but in the second regression, when the dependent variable is the log odd-ratio of choosing the lowest risk category, the flag South become positive. This kind of exercise is very useful to re-test the first regression results.

A very strong and important confirmation is taken from the variable "trading". This dummy, as said before, is an important proxy for financial knowledge. While in the first part the sign is very positive (0.62), in the second regression is negative (-

0.3582), that is coherent with our statement. People with higher financial knowledge are less risk adverse and dislike monetary (and so safe) investments.

In order to test the role of the financial knowledge and the effect of the reform for people not in the habit with financial investment, I run other regressions (R1b and R2b)[8] with the same variables of the above discussed models. In both configuration (Y=1 if the choice is for the highest risk category and Y=0 for others or Y=1 if the choice is for the lowest risk category and Y=0 for others) I introduce another variable called "Inscription after a certain date". Establishing as cut date July 2007, the coefficient of the dummy variable is negative and significant (-0.1179 with a p-value lower than 0.0001) in case of the first kind of regression and positive and significant (0.1044 with a p-value lower than 0.0001) in the second case. This further test confirms that forced participants are more risk averse. As matter of fact, this implies that employees that voluntary enrolled to the FOPEN before the reform in most of the cases made a conscious choice and suffer of lower risk aversion. While employees who joined the fund only after the reform suffer higher aversion to the risk. The important matter is that even under this specific control variable, gender continues to be fundamental in the choice. One could criticize a drawback of this variable because it is linked to the last double deep recession. In fact, it could have happen that workers were discouraged to invest in higher risk. But, it is fair to remember that data about category refers to May 2012, and moreover the variable refers to the date of inscription, that is, it is not analyzed the differences in investment behaviors before and after the cut date, but the behavior of subject enrolled before or after that date. In this case, the variable is used only to distinguish workers that joined to Fopen "voluntary" and those "forced" by the reform, and it is not used to distinguish the fund chosen before or after that date.

At this point, before going to the next model, I would like to spend few words on the goodness of fit indices of this model. As written before, one of the most important measure of goodness of fit, are those about the concordant pairs. In R1a the percentage of concordant pairs (over the maximum combination of over 150 million pairs) is 76.3%, implying a Somers'D of 0.531, a Gamma of 0.534 and a Tau-a of 0.114. It means that the model is good to shape the phenomena, assigning higher or less probability in the right direction (in the next section "the switch

[8] For the whole results see table 3 and 4.

analysis", since data are not standardized, really probability marginal effect is assigned for each observable x). In R1b, the pairs analyzed are about 108 millions, 65.9% of them are concordant. In this case the Somers' D (0.33), the Gamma (0.334) and the Tau-a (0.051) are not as good as in R1a, but still relevant. Looking at R2a and R2b, it is noticeable that the percentage of concordant and respective indices are similar to R1a R1b, showing good fit of the model. Moreover, in each regression, all the tests of joint null hypothesis (Likelihood Ratio, Score and Wald) are rejected, so restricted and unrestricted models are not equivalent.

Section 4.3 The " Switch Analysis"

In the second part of the econometric session, I take data from the switch dataset. In this case, I am able to explore in the sample more than 35 thousand switches. Switches are divided into two groups: toward riskier categories, for example from totally monetary fund to balanced bond fund, and toward less risky fund, for example from mostly equity to monetary fund. So, Y will take the value of 1 for riskier switch and 0 for safer switch.

I estimate a logistic model where the main explanatory variables are as above: gender, attitude to trade, age, wealth (cumulated balance), geographical factor and income. Income is the most controversial variable. In fact, the deep dataset used for this analysis has the drawback that data describing income refers to 2012. For this reason, it does not create a problem in the previous econometric session that refers only to 2012, but in the case of the switch analysis, such a switch could have happen since 2003 but the worker wage refers to 2012. To partially overcome this issue and to make income comparables, I divide the income of each observation by $1{,}03\text{^}(2012 - year\ of\ change)$. In this way, probably I do not obtain the true wage for that worker in that year but at least I will have the real wage of the future true income[9]. Anyway as it is possible to see in the tables, the models with or without income variable do not differ so much.

[9] It is the discounted income that the worker will obtain in the near future (and she/he can forecast quite accurately).

Looking at the most appropriate model (we call it R3a), with the most adequate and significant factors, proceeding as for the first session of tests, we can affirm the following (see table 5 including partial effects as well): the financial knowledge (flag trading), age, income and geographical factor are significant. In detail, flag for trading is not only significant but it has also an high influence on the probability to switch toward higher categories of risk. In this session, as anticipated before, since data are not standardized it is possible to calculate the average marginal effect. The logit model is non-linear function. Therefore, the parameter β_l associated with the l-th element in x_i does not directly measure the marginal effect $\frac{\partial E(y_i|x_i)}{\partial x_{il}} = \partial\pi_i/\partial x_{il}$. Rather applying the chain rule of differentiation, we obtain the marginal probability effect (MPE).

$$MPE_{il} = g(x_i^{'}\beta)\beta_l$$

The simplest method is computing the partial effects for the average individual in the sample, that is $\partial\pi_i/\partial x_{il} = \pi_i(1-\pi_i)\beta_l$, substituting $\pi_i = \bar{y}$, yields a crude estimated marginal effect of $\bar{y}(1-\bar{y})\overline{\beta_l}$. This is the method to which it is referred in this test. But I try, also to calculate the so called "average marginal probability effect" (AMPE). I include also this method because when the explanatory variables are binary, computing the effect of an infinitesimal change of x_{il} can be highly inaccurate. So to assess the average of the marginal effect of each independent variable I proceed using: $\Delta\pi_{il} = G\left(x_i^{'}\beta + \Delta x_{il}\beta_l\right) - G(x_i^{'}\beta)$. In the tables also this kind of marginal effect is presented, but it does not differ very much from the outcome obtained using the simplest method.

In this test, to be a trader increases the probability to switch for higher risk category by 32.75%. Also dummies for age affect the probability to switch to higher or lower risk categories. For example, workers younger than 36 years old have 4.44% more chance to switch to riskier sub-funds. While an individual between 46 years old and 53 years old have 2.31% less probability to pass to a higher category of risk. Looking at income, there are two significant categories (the 2 highest income level), from 2500€ to 2900€ and over 2900€ that increase probability to pass to higher category respectively by 5.43% and by 4.25%. The geographical flags work extremely well. In fact, belonging to less developed areas decrease the probability to choose higher risk categories by 4.58% if resident in the South, while by 4.73% if resident in the Islands. A resident in the North-East, the most developed area, has 3.54% more chance to switch to higher risk sub-fund. Results are in line with previous session, but in this case the variable gender is not selected by the step-wise selection method. The only mismatched is about the dummy for older than 53 years old, that enters in the regression but with an unexpected sign (even though in the complete model R3b, the sign becomes coherent with the whole study and theory).

To go deep with the results, I run a complete model (R3b) with the whole range of categories for each variable. Looking at table 6, the whole set of variables is available. Each category responds according the theoretical expectations. The only variable that does not work exactly as it does in the first econometric section (4.2) nor as in theory is the income variable. The explanation could be found out in the definition of the variable. In fact, it refers to the income of 2012 and despite the treatment (as written before), it is impossible to attribute the right wage as collected at the time of the switch. In addition, as in the stepwise regression, gender is not significant. Both for R3a and R3b the goodness of fit is excellent. For example in R3b the percentage of concordant is very high (76.4%) and in R3b it is 71.1%. Moreover, the test of the null hypothesis is rejected for the whole 3 methods taken into consideration.

As written before, our analyses is dedicated mainly to the identification of categories of workers, that due to innate aversion to risk, are not able to choose the most appropriate pension fund in order to finance their future consumption (during pension period). For this reason, to get confirmation that gender plays a crucial role in investment decision by the dataset switch as well, I proceed with the presentation of a similar regression but with a slightly different definition of the dependent variable. It is useful to highlight that till March 2003, the whole workers enrolled before that date belonged to a unique risk category called "Unico". After that date, different risk categories has been created. Of course, everybody had to choose one of the risk categories. Who did not express any preference was enrolled by default in a medium-low risk category called "Bilanciato Obbligazionario"[10]. So I try to sterilize the regression from this kind of observation[11], obtaining the

[10] See Section 3 for more specific information.

[11] In the base regression (R3a and R3b), since the switch is toward a slightly less risky category than "Unico", I taped it as Y=0.

following results (see table 7). In this case, after the sterilization of the default choice, the model R4a shows 9 dummies belonging to 4 different variables. The whole set of dummies responds in a perfect way[12]. Maybe because in this way only conscious choices are taken into account in the analysis. After that, we have many confirmations: the relationship between likelihood to switch to riskier sub-funds and age is confirmed in a monotonic way, that is, as the agent becomes older the less is the probability[13] to change for riskier sub-funds. Also income is significant and coherent with our expectations, actually, as income increases, also the probability to switch toward riskier categories increases. Geographical factor still continue to be significant and with sign compatible to the theory: living in the South (-0.1754) makes a change toward riskier assets less likely, but living in the North-East (0.2362) or North-West (0.1271) area makes riskier choice more likely. Despite so many control variables, gender dummy is now significant and positive (0.2845). In light of this regression we can confirm our intuitions, that is, over 36 years old[14] female worker living in less developed areas with a lower income is really likely to suffer pension risk.

In table 8 it is possible to have an idea of the marginal effect of single characteristics. And looking at regression R4b, where I use the same construction but avoiding a stepwise selection and including the whole regressors, it is possible to assess the overall effect of such weak worker. In fact from the model R4b, I take as default equation (where all independent dummy variables are zero) a female worker with an age between 36 and 46 years old, earning between 1900€ and 2500€ per month and living in the Centre of Italy. If such a worker instead of being female were male, the probability to be less risk averse increases by 5.76%. The next two cases indicate the changes in the predicted probability consequent to marginal changes in other demographic factors (live in the South or be older than 53 years old). It shows that an older person (over 53 years old), ceteris paribus, is

[12] In line with regressions R1a & R2a and most of the literature.

[13] Here, I refer to coefficients, but marginal effects (having the same sign as the coefficients) are available in the table 7.

[14] Even who belong in the age range 36-46 years old results a risk adverse subject, despite at this age is still highly advisable to choose a non-monetary fund.

less likely to switch to higher risk sub-fund[15]. If the worker lives in the South, the probability to pass to riskier category diminishes by 4.04%. These are just few examples, but the reader can use the tables to find whatever effect he/she likes to investigate.

The last part of this section, as outlined in the introduction, is dedicated (for completeness) to the analysis of the simultaneous first big switch of the Fopen members[16]. Excluding workers shifting from "Unico" to "Bilanciato obbligazionario" that can be considered mostly a default choice or a neutral change (since the weight of the equity was similar), 6593 choices are taken into account. 3690 observations refer to lower risk category changes ($Y = 0$), while 2903 observations relate to switches toward higher risk categories ($Y = 1$). Regression R5a, exhibited in table 9, shows even more marked results than in the whole sample regression. Results are not only significant, but also coherent with both my expectations and those of the theory. The only variable showing contrasting direction is the flag of trading, with a coefficient of -1.9032 indicating higher risk aversion for traders. But the most central variable gender is not only positive and significant but exhibit a very high coefficient, that is, looking at the odd-ratio (2.109), gender affect the risk aversion attitude considerably. Also age shows a coherent behavior (and monotonic), as worker get old preferences to switch toward higher risk categories diminish. Another consistent result is about geographical factor: workers living in less developed areas opted for lower risk categories, while those living in the North-East show lower risk aversion. In this regression it is important to remember that the variable income has been excluded before the stepwise method in order to do not affect the whole result. Then in R5b (table 10), I run a model with the same sample as R5a (only observations from "Unico") but adding the default change (from "Unico" to "Bilanciato Obbligazionario") in the $Y = 0$. In this case the observations are 27034 and result are very similar to R5a. The only difference is

[15] In this case the probability to switch toward higher risk categories decreases by 17.35%

[16] As said before, on the 16th of June 2003 has been introduced the opportunity of choice among multiple categories (4 different level of risk: "Monetario", "Bilanciato Obbligazionario", "Bilanciato" and "Prevalentemente Azionario"). So workers had to choose a category to transfer their money from "Unico" category. Who did not make any choice was assigned by default to "Bilanciato Obbligazionario".

that since this time the factor income has been included in the stepwise, it results that workers earning more than today equivalent 2900€ prefer to switch toward riskier categories, showing lower risk aversion (coherent with my expectations).

Section 4.4 Tests of the Hypothesis

In the last econometric session it is developed a test of hypothesis as in Olivares at Al (2008). We verify the existence of differences in the weight of men and women that change their portfolios within fund categories. This allows to eliminate the notion that part of previous results of the paper are driven by the number of women and men in the sample. First of all I divided the categories of risk into 2 groups: one group called High Risk Fund (HRF) and another called Low Risk Fund (LRF). To the HRF belong "Bilanciato" and "Prevalentemente Azionario". While the others, "Classe Garanzia"," Monetario", "Unico", "Bilanciato Obbligazionario" belong to LRF[17]. Then, I test the null hypothesis expressed in the following equation:

$$H_0: \left(\theta_{HRF}^{Male} - \theta_{HRF}^{Female}\right) = 0$$

That is, I verify if proportion within females moving from LRF to HRF is higher than the same motion but within men[18].

Consistent with previous findings and Olivares et al (2008), results show that both gender move from lower risky fund when investors increase their income[19]and are younger[20]. Statistically speaking (test A1), as noticeable in table 11, the difference

[17] The null hypothesis is rejected even if I take out from the sample the default option, that is, deleting from the sample workers passing from "Unico" to "Bilanciato Obbligazionario" (as in R4a and R4b). For this, see the following tests B1 and B2.

[18] It is important to distinguish this method of classification with that used in section 4.3. In fact, in this case, changes toward riskier categories but within the same group (LRF or HRF) are not included in the numerator but in the denominator of θ_{HRF} .

[19] In Olivares (2008) it is referred to wealth instead of income.

[20] Only about men

between the proportion of men that moves from LRF to HRF (21.19%) and the proportion of women doing the same (18.27%) is positive and significant (2.92%, with a p-value of the difference in weights lower than 0.0001). After introducing a filter selecting only workers with a wage higher than 2900€ the women risky weight increases slightly from 18.27% to 18.79%, while for men the weight goes from 21.19% to 21.85% (here the difference, 3.06%, is significant with a p-value of 0.005). Adding the filter age the previous filter the proportion of women decreases to 17.19%, while the weight of men continues to grow reaching 22.04%. As Olivares e Al (2008) explains, women may chose the maternity and later children education, at this age, tending to stay in lower risky funds during motherhood period. Even in this case the difference in proportions is significant. So, from this test I can affirm that women are more risk averse than men when picking pension fund portfolios.

In test B1 and B2 (table 12 and 13), I run the same test but excluding from the sample, the already illustrated, "default choice", workers switching from "Unico" to "Bilanciato Obbligazionario". In B1 it is analyzed the whole sample of subjects, while in B2 only switches done by members before 2007 are taken into account. Also in these tests, differences between men and women are significant, even after age and income filters.

Moreover, taking inspiration from this kind of test, I run another test of null hypothesis (test C1) to assess the risk aversion of forced participants. That is, as written before, most of workers, since January 2007, had to allocate their TFR in a pension fund. I test the following null hypothesis:

$$H_0: (\theta_{HRF}^{pre2007} - \theta_{HRF}^{post2007}) = 0$$

Workers that joined Fopen before 2007, did it voluntary, but those who adhered after 2007 were mainly forced. From the analysis[21], the null hypothesis is rejected and, as expected, the weight of workers passing from LRF to HRF who joined Fopen before 2007 (48.59%) is significantly higher than who adhered the fund after 2007 (38.64%). So, I may conjecture that the reform created disparity between workers comfortable with financial decision and those without "financial literacy".

[21] Results available in table 14.

Section 5 Conclusions

The aim of this paper is to understand the consequences of sensible public reform, such as pension reform in Italy and the effect on certain categories of individuals. The Italian reform on TFR ("Trattamento di fine Rapporto") provoked a very important shock in the pension system (for both households and firms). In literature, many studies used household finance data in order to model the risk aversion from an empirical point of view. Our conclusion are in line with most of the literature, but since the sample is at the same time very large and coming from effective real financial choices, this contribution to the household finance literature is quite unique. Throughout a delicate management of 3 datasets and a multitude of logistic regression (here only 10 are presented) and tests of hypothesis, I can provide significant statements on the main socio-demographic factors affecting risk aversion. Moreover, it is possible to extrapolate useful suggestions to the Italian legislator in order to reduce the pension risk.

First of all, empirical evidence has been provided in evaluating gender differences in the choice of risk in private pension funds. Not only women show higher risk aversion while choosing their most suitable category of risk, but even in the following years they tend to switch toward less risky sub-funds. The tests run in this paper highlight also higher risk aversion for elders and people living in less developed areas such as South and Islands, while higher income makes the worker less risk averse (in both relative and absolute form). Also "financial literacy" matters, in fact workers trading more are less risk averse. In addition, forced participants show higher risk aversion.

Further work could include the analysis of more sensitive information, collecting deeper information of each member through the FOPEN system. Moreover, the analysis can go beyond the usual geographical specification and could use real economic data (especially after the current double deep recession) for each city of living, providing careful statements on the effect of previous economic crisis on the individual risk aversion.

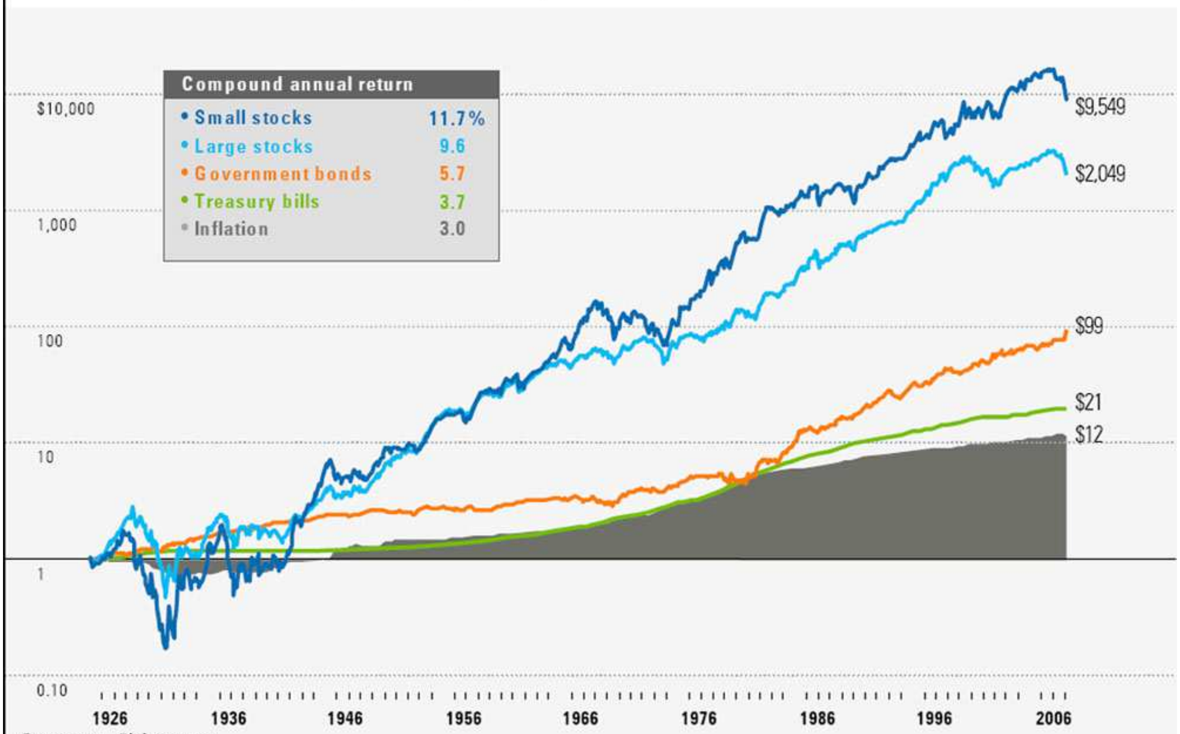
In light of these proven statements, the reform about the new destination of the TFR, issued to solve the problem of the conversion of the system to a "pay-as-you-go" programme, still have important drawbacks. As marked by other eminent studies like Cesaratto, 2008), social security risk is one of the main issues of our society, and in the future is likely to be even more significant. The TFR is roughly equal to one month's salary per year and is also called "deferred wage". It means that the TFR stocks accumulated represents the second pillar of the pension flow and it will provide a complementary pension flow fluctuating between 15% and 35% of the whole pension. In the light of these raw forecasts, the choice of risk category is fundamental, since different levels of risk correspond (theoretically and empirically) to different rate of return, as shown in figure 2 and figure 16.

As relief, the State should provide incentives (like further tax breaks applied to riskier funds) for certain categories of workers that suffer largely of higher risk aversion. Otherwise, a solution could be a more efficient educational campaign, focused on sensitizing workers on the choice of the most appropriate destination of her/his TFR.

Figures

Figure 2

Ibbotson Chart: Stocks, Bonds, Bills, & Inflation 1926-2008



Source: Ibbotson

Figure 3

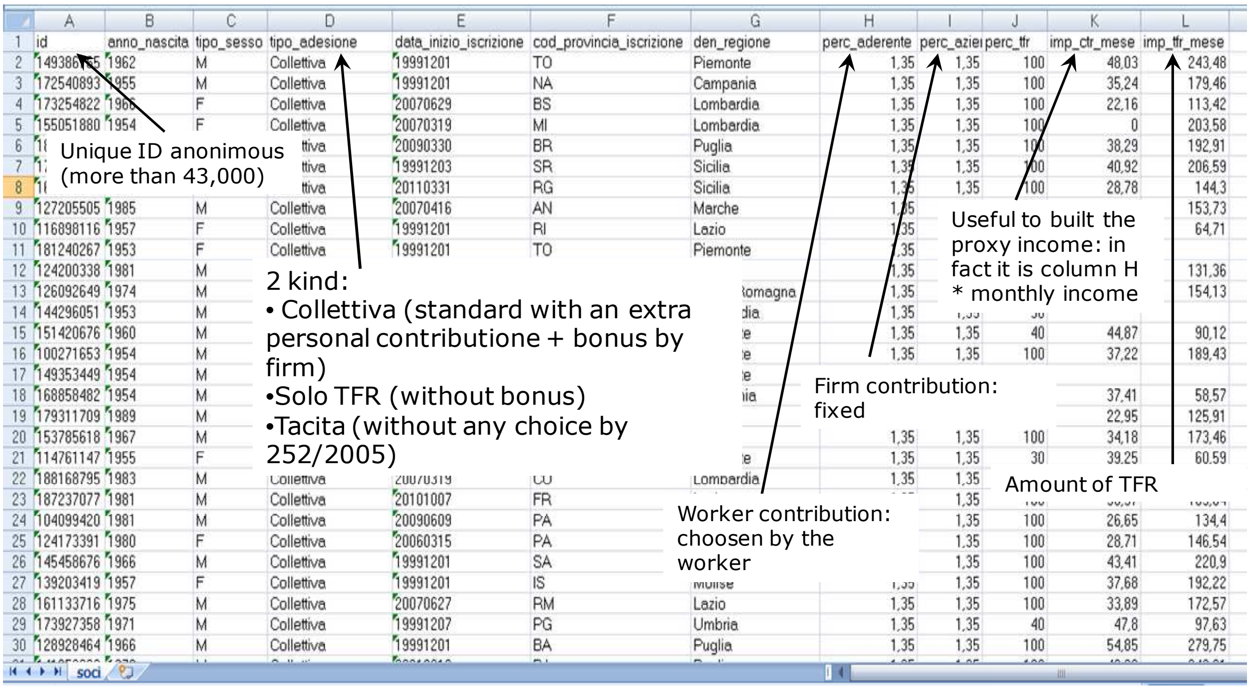


Figure 4

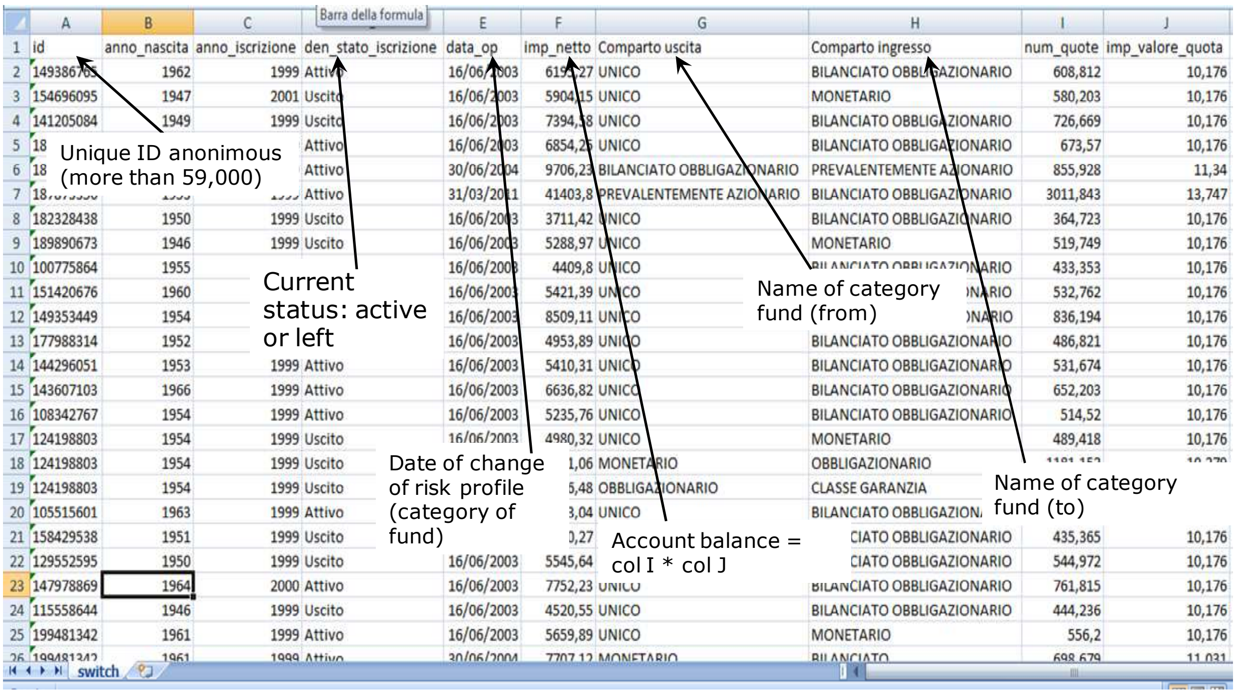
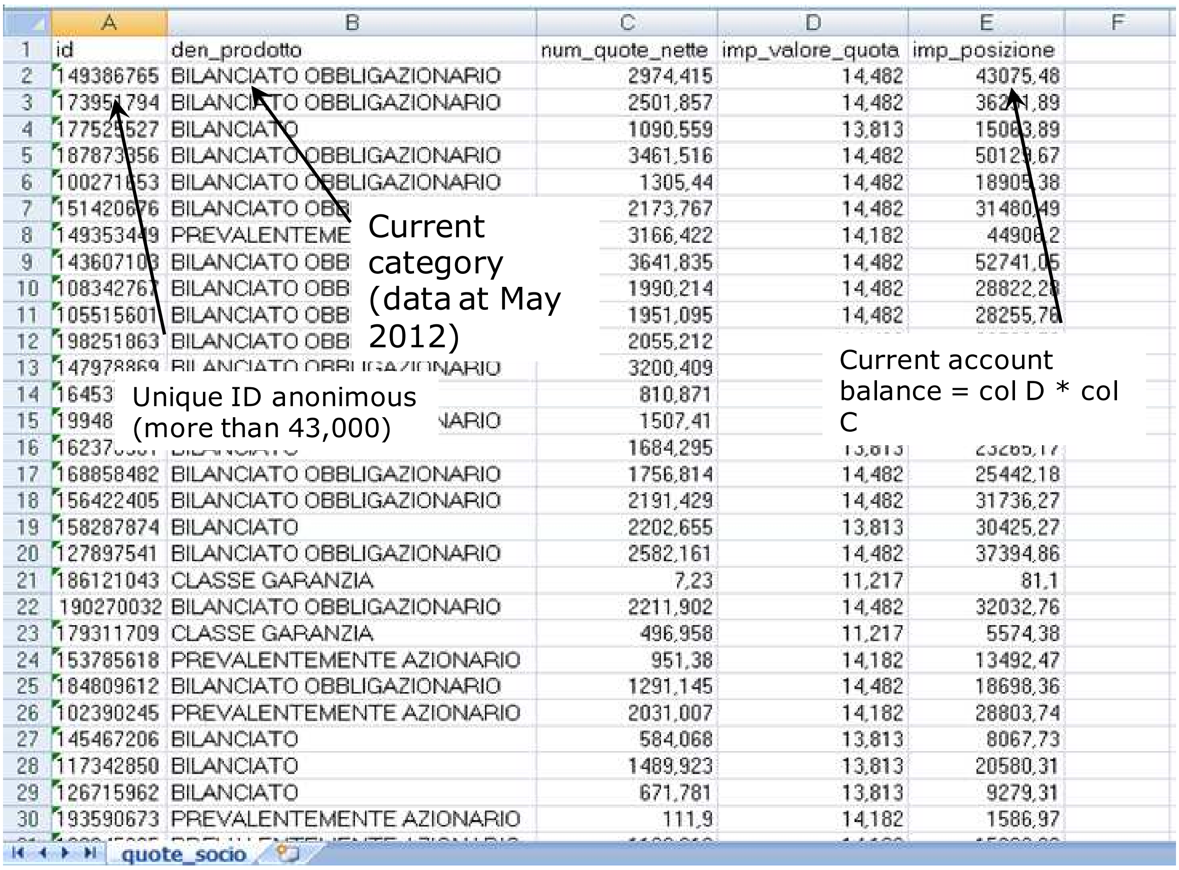


Figure 5



	A	B	C	D	E	F
1	id	den_prodotto	num_quote_nette	imp_valore_quota	imp_posizione	
2	149386765	BILANCIATO OBBLIGAZIONARIO	2974,415	14,482	43075,48	
3	173951794	BILANCIATO OBBLIGAZIONARIO	2501,857	14,482	36231,89	
4	177525527	BILANCIATO	1090,559	13,813	15063,89	
5	187873356	BILANCIATO OBBLIGAZIONARIO	3461,516	14,482	50129,67	
6	100271653	BILANCIATO OBBLIGAZIONARIO	1305,44	14,482	18905,38	
7	151420676	BILANCIATO OBB[illegible]	2173,767	14,482	31480,49	
8	149353449	PREVALENTEME[illegible]	3166,422	14,182	44906,2	
9	143607103	BILANCIATO OBB[illegible]	3641,835	14,482	52741,05	
10	108342767	BILANCIATO OBB[illegible]	1990,214	14,482	28822,28	
11	105515601	BILANCIATO OBB[illegible]	1951,095	14,482	28255,76	
12	198251863	BILANCIATO OBB[illegible]	2055,212	[illegible]	[illegible]	
13	147978869	BILANCIATO OBBLIGAZIONARIO	3200,409	[illegible]	[illegible]	
14	16453[illegible]	[illegible]	810,871	[illegible]	[illegible]	
15	19948[illegible]	[illegible]NARIO	1507,41	[illegible]	[illegible]	
16	16237[illegible]	BILANCIATO	1684,295	13,813	23265,17	
17	168858482	BILANCIATO OBBLIGAZIONARIO	1756,814	14,482	25442,18	
18	156422405	BILANCIATO OBBLIGAZIONARIO	2191,429	14,482	31736,27	
19	158287874	BILANCIATO	2202,655	13,813	30425,27	
20	127897541	BILANCIATO OBBLIGAZIONARIO	2582,161	14,482	37394,86	
21	186121043	CLASSE GARANZIA	7,23	11,217	81,1	
22	190270032	BILANCIATO OBBLIGAZIONARIO	2211,902	14,482	32032,76	
23	179311709	CLASSE GARANZIA	496,958	11,217	5574,38	
24	153785618	PREVALENTEMENTE AZIONARIO	951,38	14,182	13492,47	
25	184809612	BILANCIATO OBBLIGAZIONARIO	1291,145	14,482	18698,36	
26	102390245	PREVALENTEMENTE AZIONARIO	2031,007	14,182	28803,74	
27	145467206	BILANCIATO	584,068	13,813	8067,73	
28	117342850	BILANCIATO	1489,923	13,813	20580,31	
29	126715962	BILANCIATO	671,781	13,813	9279,31	
30	193590673	PREVALENTEMENTE AZIONARIO	111,9	14,182	1586,97	

Figure 6

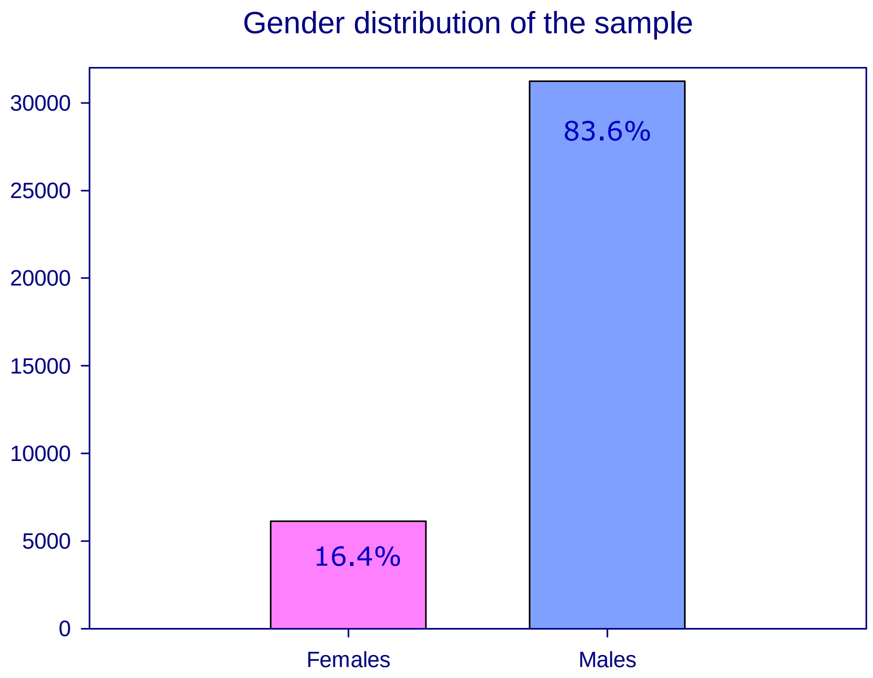


Figure 7

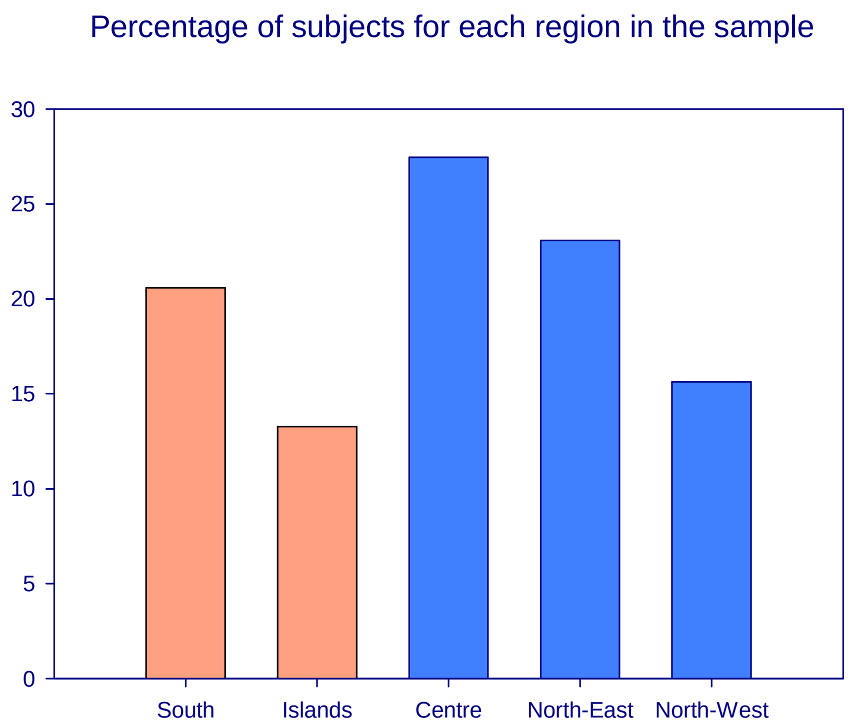


Figure 8

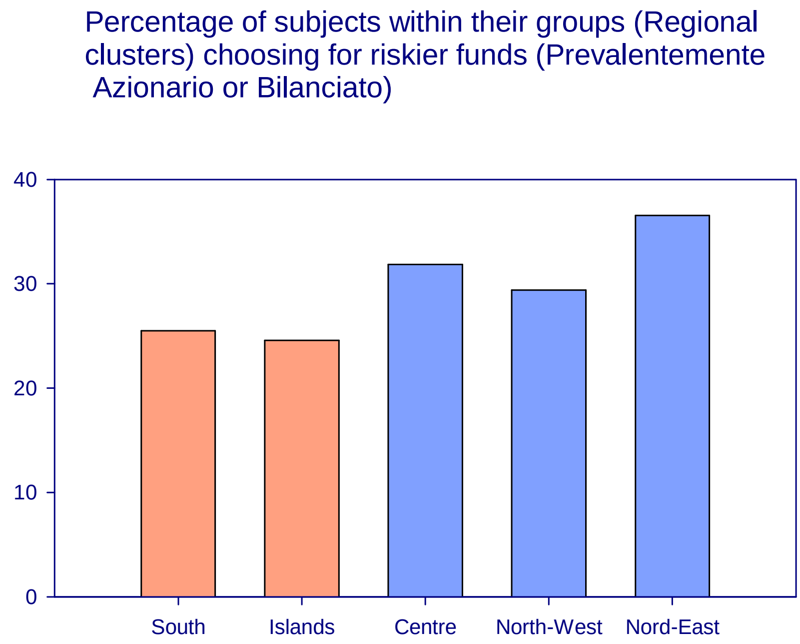


Figure 9

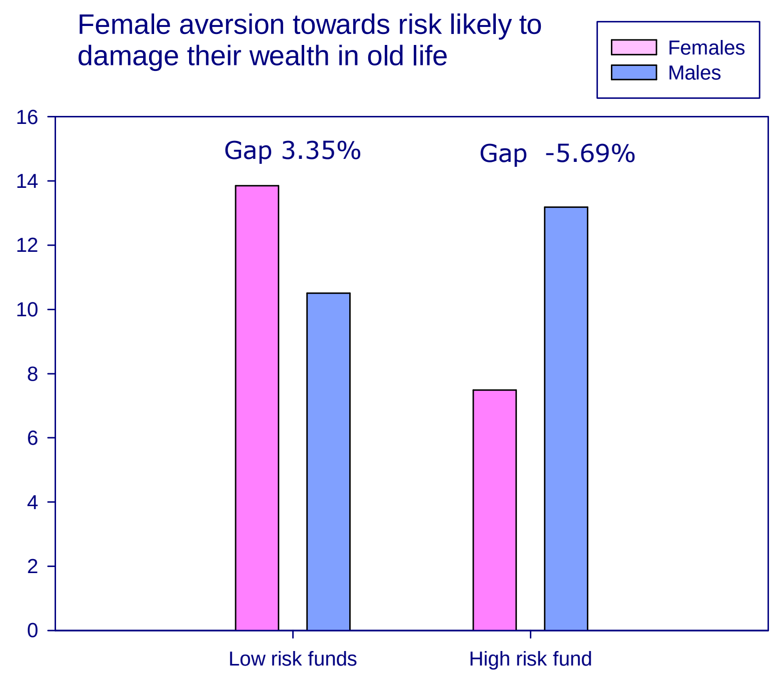


Figure 10

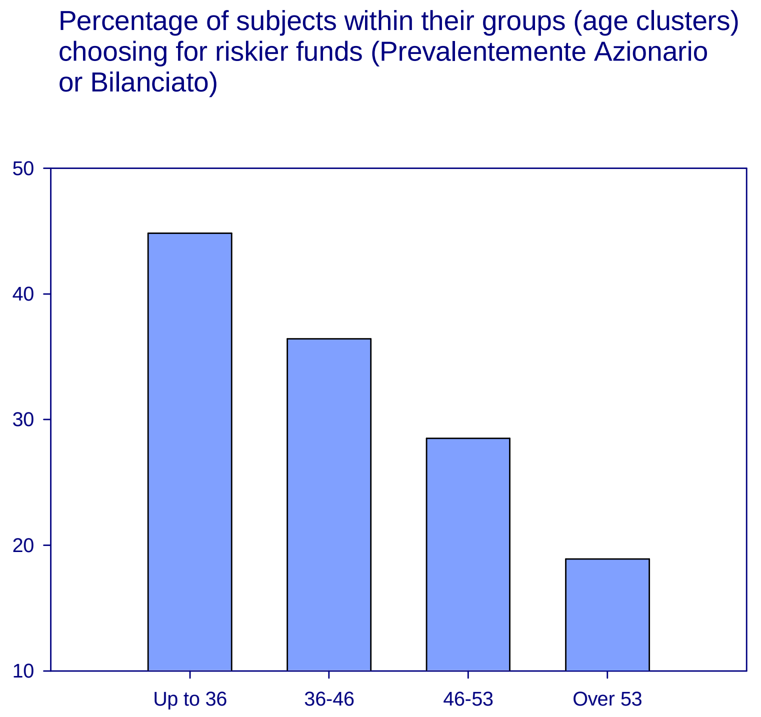


Figure 11

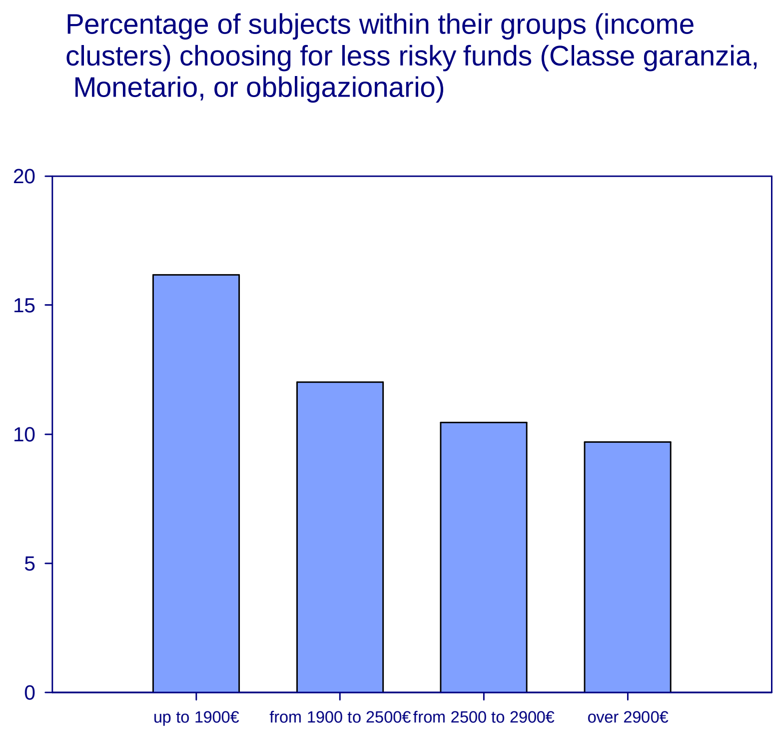


Figure 12

Figure 13

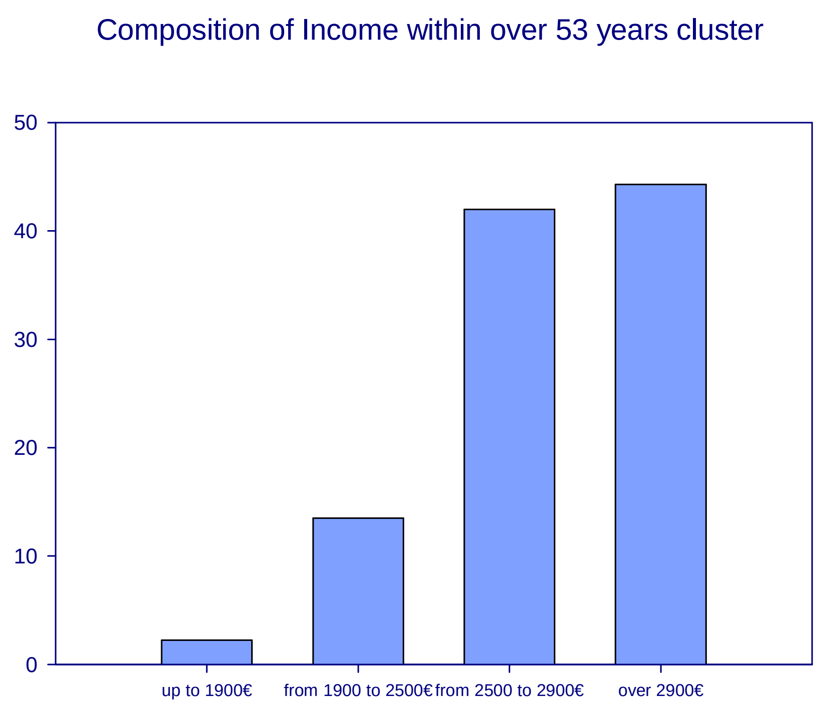


Figure 14

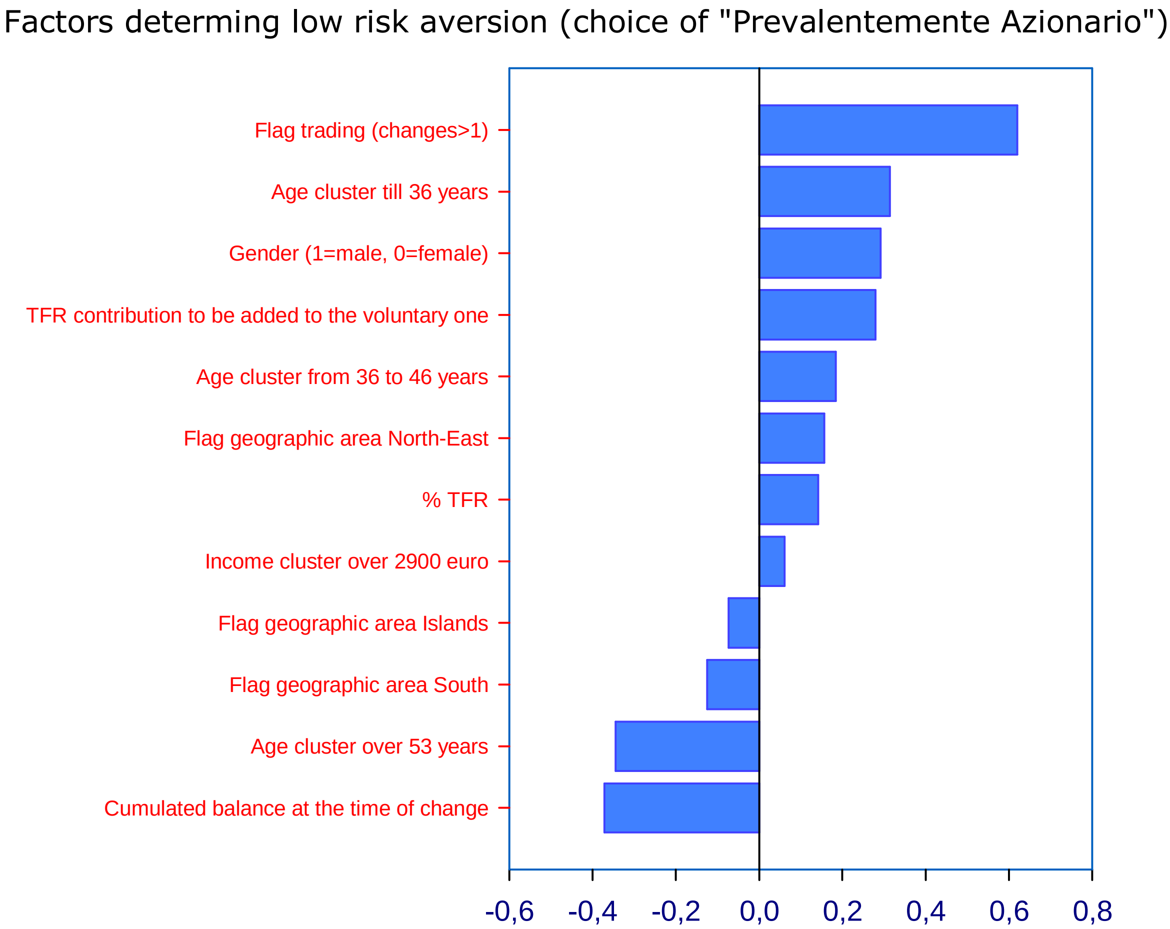
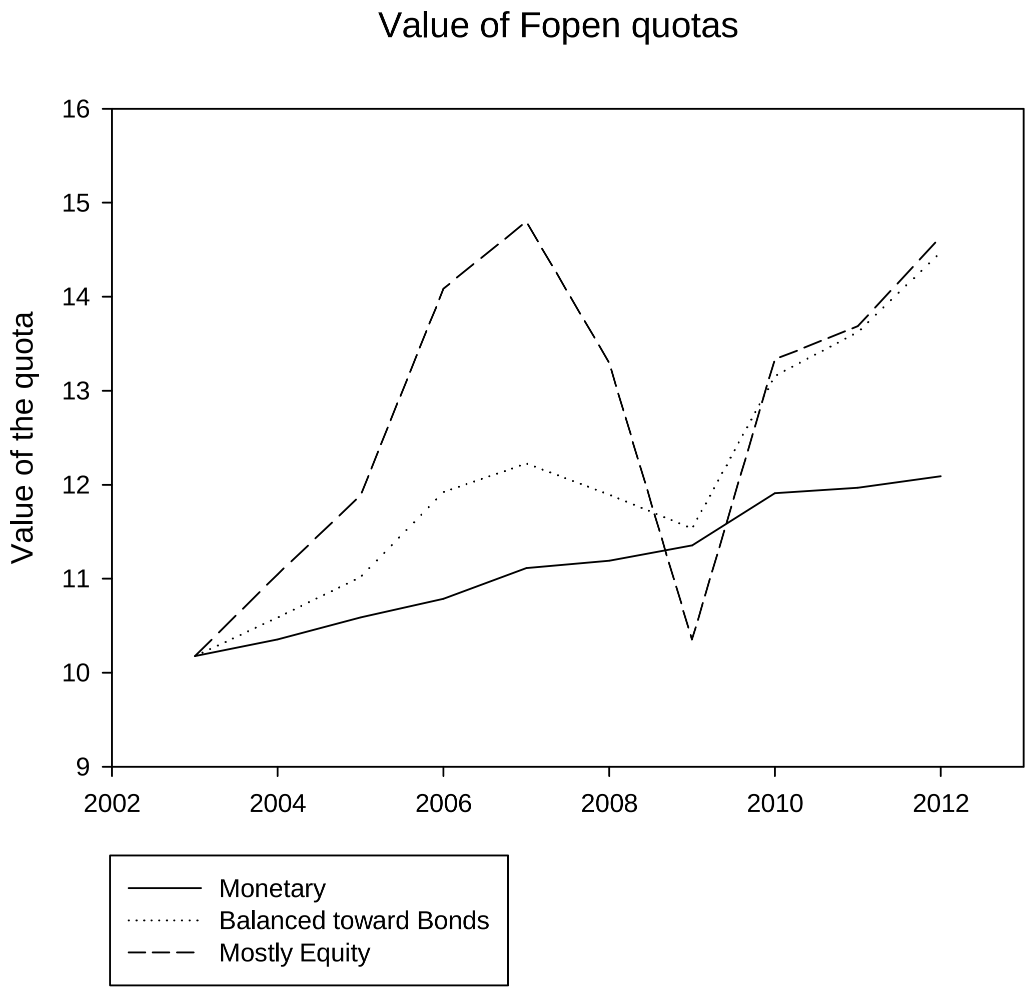

Factors determing low risk aversion (choice of "Prevalentemente Azionario")
Flag trading (changes>1)
Age cluster till 36 years
Gender (1=male, 0=female)
TFR contribution to be added to the voluntary one
Age cluster from 36 to 46 years
Flag geographic area North-East
% TFR
Income cluster over 2900 euro
Flag geographic area Islands
Flag geographic area South
Age cluster over 53 years
Cumulated balance at the time of change
-0,6
-0,4
-0,2
0,0
0,2
0,4
0,6
0,8

Figure 15

Factors determing high risk aversion (choice of "Monetario")

Figure 16


Value of Fopen quotas
Value of the quota
16
15
14
13
12
11
10
9
2002
2004
2006
2008
2010
2012
Monetary
Balanced toward Bonds
Mostly Equity

Tables

Table 1

Dependent Variable : Y=1 if worker belongs to "Azionario"; Y=0 otherwise		**R1a**			
Method: Logistic Regression					
Number of observations: 37366					
	Coefficient	Std. Error	Wald Chi-Square	Prob.	Odd Ratio
Intercept	-2.3393	0.0211	12322.2	<.0001	
Gender (1=male, 0=female)	0.2921	0.0199	214.3	<.0001	1.339
Flag trading (changes>1)	0.6202	0.0148	1755.4	<.0001	1.859
Age cluster till 36 years	0.3139	0.022	202.9	<.0001	1.369
Age cluster from 36 to 46 years	0.1838	0.0194	89.6	<.0001	1.202
Age cluster over 53 years	-0.345	0.0243	200.8	<.0001	0.708
TFR contribution	0.2794	0.0325	73.7	<.0001	1.322
% TFR	0.1417	0.031	20.9	<.0001	1.152
Cumulated balance	-0.3717	0.0258	208.2	<.0001	0.69
Income cluster over 2900 euro	0.0612	0.0235	6.78	0.0092	1.063
Flag geographic area South	-0.1246	0.0192	42.2	<.0001	0.883
Flag geographic area Islands	-0.0736	0.0187	15.5	<.0001	0.929
Flag geographic area North-East	0.1567	0.0159	96.8	<.0001	1.17

Goodness of fit through probability association				
Percentage of concordant pairs	76.3%		Somers'D	0.531
Percentage of tied pairs	0.5%		Gamma	0.534
Percentage of discordant pairs	23.2%		Tau-a	0.114
Number of pairs	150075253		c	0.766

Testing Global Null Hypothesis: BETA=0	Chi-Square	Pr > ChiSq
Likelihood Ratio	3616.3	<.0001
Score	3581.3	<.0001
Wald	2975.7	<.0001

Model Fit Statistics	Intercept only	Intercept and Covariates
AIC	27791.8	24199.5
SC	27800.3	24310.4
-2 log L	27789.8	24173.5

Table 2

Dependent Variable : Y=1 if worker belongs to "Monetario"; Y=0 otherwise		**R2a**			
Method: Logistic Regression					
Number of observations: 37366					
	Coefficient	Std. Error	Wald Chi-Square	Prob.	Odd Ratio
Intercept	-2.953	0.0256	13260.4	<.0001	
Gender (1=male, 0=female)	-0.0882	0.0212	17.3	<.0001	0.916
Flag trading (changes>1)	-0.3582	0.032	125.5	<.0001	0.699
Age cluster under 36 years old	-0.2536	0.0332	58.2	<.0001	0.776
Age cluster from 36 years to 46 years	-0.0832	0.0289	8.3	0.004	0.92
Age cluster over 53 years	0.1555	0.0267	33.9	<.0001	1.168
% TFR	-0.1988	0.0211	89.1	<.0001	0.82
Cumulated balance	-0.4102	0.0303	183.1	<.0001	0.664
Flag geographic area South	0.0821	0.0218	14.2	0.0002	1.086
Flag geographic area North-East	-0.0575	0.0247	5.4	0.02	0.944

Goodness of fit through probability association				
Percentage of concordant pairs	65.1%		Somers'D	0.321
Percentage of tied pairs	1.8%		Gamma	0.326
Percentage of discordant pairs	33.1%		Tau-a	0.035
Number of pairs	75912760		c	0.66

Testing Global Null Hypothesis: BETA=0	Chi-Square	Pr > ChiSq
Likelihood Ratio	683.6	<.0001
Score	658.2	<.0001
Wald	633.5	<.0001
Model Fit Statistics	Intercept only	Intercept and Covariates
AIC	16487.1	15821.6
SC	16495.7	15906.9
-2 log L	16485.1	15801.6

Table 3

Dependent Variable : Y=1 if worker belongs to "Azionario"; Y=0 otherwise		**R1b**			
Method: Logistic Regression					
Number of observations: 37366					
	Coefficient	Std. Error	Wald Chi-Square	Prob.	Odd Ratio
Intercept	-2.3428	0.0211	12303.5	<.0001	
Gender (1=male, 0=female)	0.2923	0.02	214.4	<.0001	1.34
Flag trading (changes>1)	0.6148	0.0148	1714.4	<.0001	1.849
Age cluster till 36 years	0.3417	0.0225	230.9	<.0001	1.407
Age cluster from 36 to 46 years	0.1838	0.0194	89.6	<.0001	1.202
Age cluster over 53 years	-0.3487	0.0244	204.8	<.0001	0.706
TFR contribution	0.3042	0.033	85.0	<.0001	1.356
% TFR	0.1461	0.0311	22.1	<.0001	1.157
Cumulated balance	-0.4513	0.0293	238.0	<.0001	0.637
Income cluster over 2900 euro	0.073	0.0236	9.5	0.002	1.076
Flag geographic area South	-0.1267	0.0192	43.6	<.0001	0.881
Flag geographic area Islands	-0.0736	0.0187	15.5	<.0001	0.929
Flag geographic area North-East	0.1541	0.016	93.3	<.0001	1.167
Inscription after 01/07/2007	-0.1179	0.02	34.68	<.0001	0.889

Goodness of fit through probability association				
Percentage of concordant pairs	76.4%		Somers'D	0.533
Percentage of tied pairs	0.5%		Gamma	0.535
Percentage of discordant pairs	23.1%		Tau-a	0.115
Number of pairs	150075253		c	0.766

Testing Global Null Hypothesis: BETA=0	Chi-Square	Pr > ChiSq
Likelihood Ratio	3651.2	<.0001
Score	3615.5	<.0001
Wald	2994.3	<.0001

Model Fit Statistics	Intercept only	Intercept and Covariates
AIC	27791.8	24199.5
SC	27800.3	24310.4
-2 log L	27789.8	24173.5

Table 4

Dependent Variable : Y=1 if worker belongs to "Monetario"; Y=0 otherwise	**R2b**				
Method: Logistic Regression					
Number of observations: 37366					
	Coefficient	Std. Error	Wald Chi-Square	Prob.	Odd Ratio
Intercept	-2.5267	0.0211	14375.4	<.0001	
Gender (1=male, 0=female)	-0.0627	0.0179	12.3	0.0005	0.93
Flag trading (changes>1)	0.264	0.0174	229.6	<.0001	1.302
Age cluster from 46 years to 53 years	0.0828	0.0247	11.2	0.0008	1.086
Age cluster over 53 years	0.2674	0.0245	118.9	<.0001	1.307
TFR contribution	0.1742	0.0325	28.8	<.0001	1.19
% TFR	-0.2686	0.0265	103.0	<.0001	0.764
Cumulated balance	-0.5805	0.0371	296.9	<.0001	0.56
Flag geographic area South	0.0471	0.0187	6.4	0.0116	1.048
Flag geographic area North-East	-0.0582	0.0204	8.1	0.0044	0.943
Inscription after 01/07/2007	0.1044	0.023	20.5	<.0001	1.11

Goodness of fit through probability association				
Percentage of concordant pairs	66.0%		Somers'D	0.331
Percentage of tied pairs	1.2%		Gamma	0.335
Percentage of discordant pairs	32.8%		Tau-a	0.051
Number of pairs	108525408		c	0.66

Testing Global Null Hypothesis: BETA=0	Chi-Square	Pr > ChiSq
Likelihood Ratio	1011.2	<.0001
Score	991.2	<.0001
Wald	950.0	<.0001
Model Fit Statistics	Intercept only	Intercept and Covariates
AIC	21725.1	20733.9
SC	21733.6	20827.7
-2 log L	21723.1	20711.9

Table 5

Dependent Variable : Y=1 if worker switches to higher risk category; Y=0 otherwise	R3a						
Method: Logistic Regression							
Number of observations: 36307							
	Coeff.	Std. Error	Wald Chi-Square	Prob.	Odd Ratio	Marginal effect Y=average	Average marginal effect
Intercept	-1.7798	0.0277	4139.5	<.0001			
Flag trading (changes>1)	1.6338	0.0260	3958.3	<.0001	5.123	0.3275	0.3205
Age cluster till 36 years	0.2214	0.0352	39.6	<.0001	1.248	0.0443	0.0392
Age cluster from 46 years to 53 years	-0.1153	0.0294	15.3	<.0001	0.891	-0.0231	-0.0198
Age cluster over 53 years	0.1941	0.0620	9.8	0.0017	1.214	0.0389	0.0347
Income cluster from 2500 to 2900	0.2707	0.0376	51.9	<.0001	1.311	0.0542	0.0485
Income cluster over 2900 euro	0.2118	0.0329	41.4	<.0001	1.236	0.0424	0.0376
Flag geographic area South	-0.2283	0.0338	45.8	<.0001	0.796	-0.0458	-0.0387
Flag geographic area Island	-0.2361	0.0382	38.2	<.0001	0.790	-0.0473	-0.0397
Flag geographic area North-East	0.1764	0.0352	25.1	<.0001	1.193	0.0353	0.0312

Goodness of fit through probability association				
Percentage of concordant pairs	71.1%		Somers'D	0.452
Percentage of tied pairs	3.0%		Gamma	0.466
Percentage of discordant pairs	25.9%		Tau-a	0.181
Number of pairs	264352012		c	0.726

Testing Global Null Hypothesis: BETA=0	Chi-Square	Pr > ChiSq
Likelihood Ratio	3616.3	<.0001
Score	3581.3	<.0001
Wald	2975.7	<.0001

Model Fit Statistics	Intercept only	Intercept and Covariates
AIC	42893.4	38026.8
SC	42901.9	38111.8
-2 log L	42891.4	38006.8

Table 6

Dependent Variable : Y=1 if worker switches to higher risk category; Y=0 otherwise	**R3b**						
Method: Logistic Regression							
Number of observations: 36307							
	Coeff.	Std. Error	Wald Chi-Square	Prob.	Odd Ratio	Marginal effect Y=average	Average marginal effect
Intercept	-2.2436	0.0466	2322.4	<.0001			
Flag trading (changes>1)	1.1560	0.0294	1547.7	<.0001	3.177	0.2318	0.2138
Gender (1=male, 0=female)	-0.0169	0.0357	0.2254	0.635	0.983	-0.0034	-0.0027
Age cluster till 36 years	0.2974	0.0361	67.8	<.0001	1.346	0.0596	0.0501
Age cluster from 36 to 46 years							
Age cluster from 46 years to 53 years	-0.1886	0.0306	37.9	<.0001	0.828	-0.0378	-0.0304
Age cluster over 53 years	-0.4998	0.0696	51.5	<.0001	0.607	-0.1002	-0.0744
Income cluster less than 1900 euro	-0.0280	0.0348	0.6	0.42	0.972	-0.0056	-0.0045
Income cluster from 1900 to 2500							
Income cluster from 2500 to 2900	0.0190	0.0409	0.2	0.643	1.019	0.0038	0.00312
Income cluster over 2900 euro	-0.2500	0.0382	42.9	<.0001	0.779	-0.0501	-0.0395
Net value	0.0001	3.141E-6	1263.6	<.0001	1.000	2.25E-5	
Flag geographic area Islands	-0.3519	0.0433	65.9	<.0001	0.703	-0.0706	-0.0550
Flag geographic area South	-0.2242	0.0387	33.7	<.0001	0.799	-0.0449	-0.0358
Flag geographic area Centre							
Flag geographic area North-West	0.0111	0.0377	0.1	0.769	1.011	0.0022	0.0018
Flag geographic area North-East	0.1963	0.0398	24.3	<.0001	1.217	0.0394	0.0329

Goodness of fit through probability association				
Percentage of concordant pairs	76.4%		Somers'D	0.533
Percentage of tied pairs	0.5%		Gamma	0.535
Percentage of discordant pairs	23.1%		Tau-a	0.115
Number of pairs	150075253		c	0.766

Testing Global Null Hypothesis: BETA=0	Chi-Square	Pr > ChiSq
Likelihood Ratio	6449.7	<.0001
Score	6553.9	<.0001
Wald	5212.8	<.0001
Model Fit Statistics	Intercept only	Intercept and Covariates
AIC	42893.4	36469.7
SC	42901.9	36588.7
-2 log L	42891.4	36441.7

The parameters "Age cluster from 36 to 46 years", "Income cluster from 1900 to 2500" and "Flag geographic area Centre" have been set to 0, since the categories are a linear combination of others.

Table 7

Dependent Variable : Y=1 if worker switches to higher risk category; Y=0 otherwise		**R4a**					
Method: Logistic Regression							
Number of observations: 15866							
	Coeff.	Std. Error	Wald Chi-Square	Prob.	Odd Ratio	Marginal effect Y=average	Average marginal effect
Intercept	0.4451	0.0564	62.1750	<.0001			
Gender (1=male, 0=female)	0.2845	0.0432	43.2631	<.0001	1.329	0.0659	0.0664
Age cluster from 36 to 46 years	-0.1155	0.0486	5.6537	0.0174	0.891	-0.026	-0.0263
Age cluster from 46 years to 53 years	-0.3528	0.0512	47.5660	<.0001	0.703	-0.081	-0.0816
Age cluster over 53 years	-0.6420	0.0745	74.2646	<.0001	0.526	-0.148	-0.1539
Income cluster from 2500 to 2900	0.1567	0.0484	10.4650	0.0012	1.170	0.0363	0.0351
Income cluster over 2900 euro	0.1815	0.0437	17.2751	<.0001	1.199	0.0420	0.0407
Flag geographic area South	-0.1754	0.0439	15.9873	<.0001	0.839	-0.0406	-0.0406
Flag geographic area North-West	0.1271	0.0456	7.7788	0.0053	1.136	0.0294	0.0287
Flag geographic area North-East	0.2362	0.0485	23.7052	<.0001	1.266	0.0547	0.0527

Goodness of fit through probability association				
Percentage of concordant pairs	55.6%		Somers'D	0.153
Percentage of tied pairs	4.1%		Gamma	0.160
Percentage of discordant pairs	40.3%		Tau-a	0.071
Number of pairs	58327173		c	0.577

Testing Global Null Hypothesis: BETA=0	Chi-Square	Pr > ChiSq
Likelihood Ratio	245.8	<.0001
Score	246.4	<.0001
Wald	242.5	<.0001

Model Fit Statistics	Intercept only	Intercept and Covariates
AIC	20821.3	20593.4
SC	20829.0	20670.2
-2 log L	20819.301	20573.4

Table 8

Dependent Variable : Y=1 if worker switches to higher risk category; Y=0 otherwise	**R4b**						
Method: Logistic Regression							
Number of observations: 15866							
	Coeff.	Std. Error	Wald Chi-Square	Prob.	Odd Ratio	Marginal effect Y=average	Average marginal effect
Intercept	0.2572	0.0592	18.9	<.0001			
Flag trading (changes>1)	-0.3086	0.0415	55.2	<.0001	0.734	-0.0715	-0.0682
Gender (1=male, 0=female)	0.2488	0.0438	32.2	<.0001	1.283	0.0576	0.0571
Age cluster till 36 years	0.1617	0.0501	10.4	0.0012	1.176	0.0374	0.0358
Age cluster from 36 to 46 years							
Age cluster from 46 years to 53 years	-0.2634	0.0392	45.2	<.0001	0.768	-0.0610	-0.0599
Age cluster over 53 years	-0.7492	0.0696	115.7	<.0001	0.473	-0.1735	-0.1773
Income cluster less than 1900 euro	-0.1064	0.0457	5.4	0.0198	0.899	-0.0246	-0.0241
Income cluster from 1900 to 2500							
Income cluster from 2500 to 2900	0.0246	0.0513	0.2	0.6311	1.025	0.0057	0.0055
Income cluster over 2900 euro	-0.0615	0.0485	1.6	0.2045	0.940	-0.0142	-0.0138
Net value	0.00004	3.128E-6	194.1	<.0001	1.000	1.02E-5	
Flag geographic area Islands	-0.0654	0.0555	1.4	0.2381	0.937	-0.0404	-0.0147
Flag geographic area South	-0.1745	0.0482	13.1	0.0003	0.840	-0.0151	-0.0398
Flag geographic area Centre							
Flag geographic area North-West	0.1089	0.0495	4.8	0.0279	1.115	0.02523	0.0242
Flag geographic area North-East	0.2323	0.0523	19.7	<.0001	1.261	0.05382	0.0512

Goodness of fit through probability association				
Percentage of concordant pairs	61.4%		Somers'D	0.234
Percentage of tied pairs	0.7%		Gamma	0.236
Percentage of discordant pairs	37.9%		Tau-a	0.109
Number of pairs	58327173		c	0.617

Testing Global Null Hypothesis: BETA=0	Chi-Square	Pr > ChiSq
Likelihood Ratio	462.2	<.0001
Score	450.6	<.0001
Wald	435.9	<.0001
Model Fit Statistics	Intercept only	Intercept and Covariates
AIC	20821.3	20385.1
SC	20829.0	20492.5
-2 log L	20819.3	20357.1

The parameters "Age cluster from 36 to 46 years", "Income cluster from 1900 to 2500" and "Flag geographic area Centre" have been set to 0, since the categories are a linear combination of others.

Table 9

Dependent Variable : Y=1 if worker switches to higher risk category; Y=0 otherwise	**R5a**				
Method: Logistic Regression					
Number of observations: 6593 (from "Unico")					
	Coeff.	Std. Error	Wald Chi-Square	Prob.	Odd Ratio
Intercept	-0.3203	0.1284	6.2	0.0126	
Flag trading (changes>1)	-1.9032	0.0626	924.5	<.0001	0.149
Gender (1=male, 0=female)	0.7464	0.0787	90.0	<.0001	2.109
Age cluster till 36 years	0.2204	0.0741	8.8	0.0029	1.247
Age cluster from 46 years to 53 years	-0.7148	0.0674	112.3	<.0001	0.489
Age cluster over 53 years	-2.8501	0.6127	21.6	<.0001	0.058
Net value	0.0001	0.00002	25.1	<.0001	1.000
Flag geographic area Islands	-0.7432	0.0895	68.9	<.0001	0.476
Flag geographic area South	-0.7022	0.1037	45.9	<.0001	0.496
Flag geographic area Centre	-0.2097	0.0784	7.2	0.0075	0.811
Flag geographic area North-East	0.3203	0.0867	13.7	0.0002	1.378

Goodness of fit through probability association				
Percentage of concordant pairs	78.0%		Somers'D	0.562
Percentage of tied pairs	0.2%		Gamma	0.564
Percentage of discordant pairs	21.8%		Tau-a	0.277
Number of pairs	10712070		c	0.781

Testing Global Null Hypothesis: BETA=0	Chi-Square	Pr > ChiSq
Likelihood Ratio	1663.0	<.0001
Score	1490.0	<.0001
Wald	1228.1	<.0001
Model Fit Statistics	Intercept only	Intercept and Covariates
AIC	9047.7	7404.6
SC	9054.5	7479.4
-2 log L	9045.7	7382.6
The variable income has been left before the stepwise method		

Table 10

Dependent Variable : Y=1 if worker switches to higher risk category; Y=0 otherwise		**R5b**			
Method: Logistic Regression					
Number of observations: 27034 (from "Unico")					
	Coeff.	Std. Error	Wald Chi-Square	Prob.	Odd Ratio
Intercept	-2.6092	0.0773	1138.6	<.0001	
Flag trading (changes>1)	-0.4495	0.0524	73.5	<.0001	0.638
Gender (1=male, 0=female)	0.3905	0.0606	41.5	<.0001	1.478
Age cluster from 36 years to 46 years	-0.1719	0.0496	12.0	0.0005	0.842
Age cluster from 46 years to 53 years	-0.7296	0.0596	149.7	<.0001	0.482
Age cluster over 53 years	-2.2800	0.5831	15.3	<.0001	0.102
Income cluster over 2900 euro	0.1478	0.0532	7.7	0.0055	1.159
Flag geographic area Centre	0.5690	0.0568	100.4	<.0001	1.767
Flag geographic area North-West	0.6807	0.0568	143.5	<.0001	1.975
Flag geographic area North-East	0.9988	0.0588	288.6	<.0001	2.715

Goodness of fit through probability association				
Percentage of concordant pairs	62.5%		Somers'D	0.289
Percentage of tied pairs	3.9%		Gamma	0.301
Percentage of discordant pairs	33.6%		Tau-a	0.055
Number of pairs	70052293		c	0.645

Testing Global Null Hypothesis: BETA=0	Chi-Square	Pr > ChiSq
Likelihood Ratio	677.8	<.0001
Score	641.1	<.0001
Wald	607.4	<.0001
Model Fit Statistics	Intercept only	Intercept and Covariates
AIC	18439.7	17779.9
SC	18447.9	17862.0
-2 log L	18437.7	17759.9

Table 11

Test of hypothesis **A1**	Male	Female	Differences in weights	Chi-Square	Pr > ChiSq
Total (34617)	21.19%	18.27%	2.92%	23.85	<0.0001
If Income > 2500€ (15174)	21.85%	18.79%	3.06%	7.83	0.0052
If Income> 2500€ & Age <37 years old (1592)	22.04%	17.19%	14.83%	3.29	0.0698
In parenthesis the size of the sample subgroup					

Table 12

Test of hypothesis **B1**	Male	Female	Differences in weights	Chi-Square	Pr > ChiSq
Total (14935)	50.28%	37.68%	12.60%	138.18	<0.0001
If Income > 2500€ (5530)	49.44%	38.43%	11.01%	27.58	<0.0001
If Income> 2500€ & Age <37 years old (480)	56.92%	45.36%	11.56%	4.17	0.0411
In parenthesis the size of the sample subgroup					

Table 13

Test of hypothesis **B2**	Male	Female	Differences in weights	Chi-Square	Pr > ChiSq
Total (14125)	50.89%	37.55%	13.34%	143.73	<0.0001
If Income > 2500€ (5343)	49.97%	39.63%	10.34%	22.71	<0.0001
If Income> 2500€ & Age <37 years old (400)	60.19%	51.85%	8.34%	1.85	0.1739
In parenthesis the size of the sample subgroup					

Table 14

Test of hypothesis **C1**	Inscription before 01/01/2007	Inscription after 01/01/2007	Differences in weights	Chi-Square	Pr > ChiSq
Total (14935)	48.59%	38.64%	9.95%	30.36	<0.0001
If Income > 2500€ (5530)	48.81%	29.95%	18.86%	25.58	<0.0001
If Income> 2500€ & Age <37 years old (480)	58.58%	35.00%	23.58%	14.85	0.0001
In parenthesis the size of the sample subgroup					

Bibliography

1. Bajtelsmit, V. L. and J.L. Vanderheid. 1997. "Risk Aversion and Pension Investment Choice." In Positioning Pension for the Twenty-First Century. Edited by M. S. Gordon, O.S. Mitchell and M.M. Twinney. University of Pennsylvania Press.
2. Bajtelsmit, V., A. Bernasek, and N. Jianakoplos. 1999. "Gender differences in defined contribution pension decisions". Financial Services Review. nos 8, 1-10.
3. Beckmann, Daniela and Menkhoff, Lukas. 2008. "Will women be women?: analyzing the gender difference among financial experts". Discussion papers School of Economics and Management of the Hanover Leibniz University, No. 391.
4. Bernasek, Alexandra and Stephanie Shwiff. 2001. "Gender, Risk and Retirement," Journal of Economic Issues, 35 (2), 345-56.
5. Cameroon A. Colin and Trivedi Pravin K. 2005. "Microeconometrics, Methods and Applications".
6. Cappelletti Giuseppe G. L., G. Guazzarotti and P. Tommasino. 2010. "The effect of age on portfolio choices: evidence from an Italian pension fund". Banca d'Italia Working Papers 768.
7. Cappelletti Giuseppe G. L., G. Guazzarotti and P. Tommasino. 2010. "Le scelte previdenziali nell'indagine sui bilanci delle famiglie della Banca d'Italia". Banca d'Italia Occasional Papers n.77.
8. Cesaratto Segio. 2008. "The Macroeconomics of the Pension Fund Reforms and the case of the severance pay reform in Italy". Quaderni del Dipartimento di Economia Politica dell'Università degli Studi di Siena.
9. Felton, J., Gibson, B., & Sanbonmatsu, D. M. 2003. "Preference for risk as a function of trait optimism and gender". Journal of behavioral finance, 4, 33-48
10. Friend, I., and M.E. Blume 1975. "The Demand for Risky Assets", American Economic Review, Vol. 65(5), pp. 900-922.
11. Fornero E., Lusardi A. e Monticone C. 2009. "Adequacy of Saving for Old Age in Europe", WP CeRP n. 87/09.
12. Grable, J. E., Lytton, R. H., & O'Neill, B. 2004. "Projection bias and financial risk tolerance". The Journal of Behavioral Finance, 5(3), 142–147.

13. Greene William H. "Econometric Analysis, Sixth Edition" .
14. Guiso L. and P. Sodini. 2012. "Household finance. An emerging field". EIEF Working Paper 04/12 .
15. Guiso, Luigi, and Monica Paiella. 2008. "Risk Aversion, Wealth and Background Risk." Journal of the European Economic Association, 6(6):1109–1150.
16. Guiso, Luigi, Paola Sapienza, and Luigi Zingales. 2011a. "Time Varying Risk Aversion" Einaudi Institute for Economics and Finance, mimeo.
17. Haliassos M. and Bertaut C. C. 1995. "Why do so few hold stocks". The Economic Journal, 105, 1110-1129.
18. Jianakoplos, N. and A. Bernasek. 1998. "Are Women More Risk Averse?". Economic Inquiry, Vol. XXXVI: 620-630.
19. Kachelmeir, Steven J., and Mohamed Shehata. 1992. "Examining Risk Preferences under High Monetary Incentives: Experimental Evidence from the People's Republic of China. "American Economic Review, 85(2): 1120-1141.
20. Kahneman, Daniel, and Amos Tversky. 1979. "Prospect Theory: An Analsysis of Choice Under Risk." Econometrica, 47(2): 263-291.
21. Olivares Josè A., D. Diaz, M. Besser. 2008. "Gender and Portfolio Choice: Are Women More Risk Averse when Selecting Pension Funds?".
22. Settimo C. and Cannata R. 2007. "Italian Complementary Pension Reform". Assicurazioni Generali.
23. Skinner J. 2007. "Are You Sure You're Saving Enough for Retirement?", The Journal of Economic Perspectives, Vol. 21, No. 3 (Summer, 2007), pp. 59-80.
24. Sundén, A. and B. Surette. 1998. "Gender Differences in the Allocation of Assets in Retirement Savings Plans". The American Economic Review, Vol. 88, N°2: 207-211.
25. Winkelmann Rainer and Boes Stefan. 2006. "Analysis of Microdata".