Financial Internet Quarterly 2024, vol. 20 / no. 2, p. 1-14
10.2478/fiqf-2024-0008

MODELING FINANCIAL RISK ATTITUDE: THE ROLE OF EDUCATION AND FINANCIAL LITERACY

Maria Iannario[1], Anna Clara Monti[2], Domenico Scalera[3]

Abstract

This paper studies the relationship between risk propensity, education and financial literacy. The results of the empirical investigation confirm the importance of the key explanatory variables of education and financial competence. Since they are both included in the model, the different roles of each are singled out. In particular, while education turns out to be a factor contributing to raising risk tolerance, financial literacy tends to reduce risk propensity. Risk attitude is evaluated by self-reported assessment and modeled through cumulative logit models. In order to handle anomalous data, M estimators with a bounded influence function are considered.

JEL classification: C13, D14, G53

Keywords: Financial Literacy, Ordinal Response Models, Risk Attitude, Robust Estimation

Received: 08.02.2024 Accepted: 05.04.2024

Cite this:

Iannario M., Monti Clara A. & Scalera D. (2024). Modeling financial risk attitude: the role of education and financial literacy. Financial Internet Quarterly 20(2), pp. 1-14.

[1] University of Naples Federico II, Department of Political Science, Italy, e-mail: maria.iannario@unina.it, https://orcid.org/0000-0002-2646-9937.
[2] University of Sannio, Department of Law and Economics, Italy, e-mail: acmonti@unisannio.it, https://orcid.org/0000-0003-3369-884X.
[3] University of Sannio, Department of Law and Economics, Italy, e-mail: scalera@unisannio.it, https://orcid.org/0000-0001-5582-6924.

Introduction

The determinants of the risk propensity/aversion of economic agents have long drawn the interest of scholars, at least since the studies of Cohn et al. (1975) as well as Friend and Bleme (1975), because of the crucial implications for economic theory and policy making, the functioning of financial markets, portfolio choices and demand for insurance.

We consider models for self-reported risk propensity, where survey respondents are asked to classify themselves into ordered categories of risk attitude. Whether survey questions are an appropriate source to derive measures of individual risk aversion is a controversial issue. Nevertheless, many researchers have used self-assessed risk propensity (e. g. Courbage et al., 2018; Fisher & Yao, 2017; Yao et al., 2011). Although lack of incentive compatibility, inattention or strategic motives may in principle induce respondents not to reveal their true preferences, the behavioral validity of survey responses has been shown by Dohmen et al. (2011).

Self-reported risk attitude is typically collected as an ordinal response. Unlike previous studies (e.g. Courbage et al., 2018) that collapse categories to apply binary logistic regression, we propose modeling self-reported risk attitude through a cumulative model (McCullagh, 1980) which takes into account the ordinal nature of the response and is more parsimonious than alternative models used in this context such as the continuation-ratio (e.g. Fisher & Yao, 2017; Yao et al., 2011).

Ordered response models are usually fitted through Maximum Likelihood Estimators (MLE). However anomalous data, produced by extreme design points and/or incoherent responses can negatively affect the reliability of likelihood-based estimation and testing. To cope with the occurrence of anomalous data, M estimators are proposed. These estimators have a bounded influence function and therefore yield sound inferential results in case of outliers.

The model is applied to the risk attitude of Italian Households recorded by the Bank of Italy Survey of Household Income and Wealth (SHIW). In accordance with the literature, we focus on the hypothesis that risk propensity depends on education and financial literacy, along with a number of other controls. Actually, while many personal (De Paola, 2012; Fisher & Yao, 2017), social (Doepke & Zilibotti, 2005) and economic (Guiso & Paiella, 2008) factors, such as age, gender, religion, income, wealth, health, parents' attitude, have been considered among the determinants of risk tolerance, more recently researchers have been especially concerned with the role of education and financial literacy. Although most contributions seem to be inclined to state a positive relationship between risk propensity and education (or financial literacy), the issue remains basically unsettled (Outreville, 2015).

The earliest studies by Hersch (1996) as well as Jianakoplos and Bernasek (1998) document a negative relationship between risk propensity and education. In the same vein, considering that more educated people usually have a higher demand for insurance, Browne et al. (2000), Hwang and Gao (2003) as well as Jung (2015) reach a similar conclusion. Conversely, Bellante and Green (2004), analyzing the behavior of a large sample of individuals in the USA, compare college and high school graduates with people without a diploma, to show that more educated individuals have greater financial risk propensity. Similar results are found by Chong and Martinez (2019) for Peru et al. (2007) for Denmark, Kapteyn and Teppa (2011) in a sample of Dutch households, and Lin (2009) for Taiwan. Other authors highlight that the relationship might be nonlinear (Barsky et al., 1997) or simply not statistically significant (Halek & Eisenhauer, 2001).

A specific branch of the literature focuses on the impact of financial literacy on risk aversion, recognizing a specific role for financial literacy which "requires additional investment not currently part of a general education" (Lusardi & Mitchell, 2023, p. 142). Bayer et al. (2009) show that better-informed people become more risk averse and save a larger share of income. Similarly, van Rooij et al. (2011a), working on a survey of Dutch households, find that those with better financial knowledge are more likely to plan for retirement. More recently, Sutter et al. (2020) maintain that financial literacy makes subjects more risk averse. On the other hand, most of the literature reaches opposite conclusions. Van Rooij et al. (2011b) observe that higher financial literacy leads to higher propensity to risk and investments in stock. A positive association between financial literacy and risk tolerance is also detected by Bannier and Neubert (2016); Bayar et al. (2020) and Li et al. (2020). Dimmock et al. (2016) confirm this result but also find a negative association between risk tolerance and high-school education. According to Stoian et al. (2021), risk propensity increases with financial literacy if the former is measured with an experimental method but not through self-assessment. Finally, Hermansson and Johnson (2021) interestingly argue that more than financial literacy, what really matters in shaping risk tolerance is financial interest (i.e. the intrinsic motivation and attention of an agent toward financial information), which would drive both literacy and attitude toward risk.

Having in mind this background, this paper studies the relationship between risk propensity, education and financial literacy. Robust M estimation allows us to properly assess the relevance of the covariates. The results confirm the importance of the key explanatory variables: overall education and financial competence. Since they are both included in our model, the different roles of each of them can be singled out. In particular, while education turns out to be a factor contributing to raising risk tolerance, financial literacy tends to reduce risk propensity. Finally, M estimators recognize relevance to the income – financial literacy interaction terms, which reinforce the role of the latter variable for more affluent households.

The remainder of the paper is organized as follows. The second section illustrates the dataset, providing details on the construction of the key variables of risk attitude, education, and financial literacy. The third and fourth sections respectively deal with the statistical models and introduce robust estimation methods. The empirical results are in the fifth section, while final remarks end the paper. The Appendix collects some technical details.

Data and variables

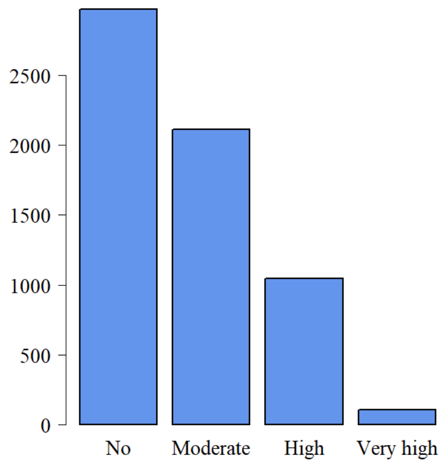
Our investigation hinges on a rich dataset provided by the 2020 issue of the Bank of Italy Survey of Household Income and Wealth (SHIW) (The SHIW survey has gathered data on the income and savings of Italian households since the 1960's, growing in scope over time, and now including several aspects of households' economic and financial behavior. The sample comprises about 7,000 households over about 300 Italian municipalities. The results are regularly published at the website https://www.bancaditalia.it/pubblicazioni/indagini-famiglie/index.html, gathering data on Italian households. Self-reported risk propensity is retrieved through the question: 'In managing your financial investments, would you say to have a preference for investments that offer (...)' with answers:

1) low return with no risk of losing invested capital;
2) fair return with good degree of protection for invested capital;
3) high return with fair degree of protection for the invested capital;
4) very high return with high risk of losing invested capital.

The answers 1 (no risk), 2 (moderate risk), 3 (high risk), and 4 (very high risk) correspond to the categories of the response variable Y. The unconditional distribution of risk attitude of our sample of 6239 households, depicted in Figure 1, shows that a high share of individuals is risk averse, while fewer and fewer individuals display higher risk propensity.

Regarding the explanatory variables, the overall educational level is measured by the two binary variables Secondary and Tertiary, taking value 1 when the respondent has a secondary school degree or a university (or higher) degree, respectively, and 0 otherwise (the baseline category being educational levels lower than the secondary school). Unlike most of the literature, we proxy overall education with both high-school and university degrees to get a finer assessment of the role of education.

Figure 1: Distribution of risk attitude

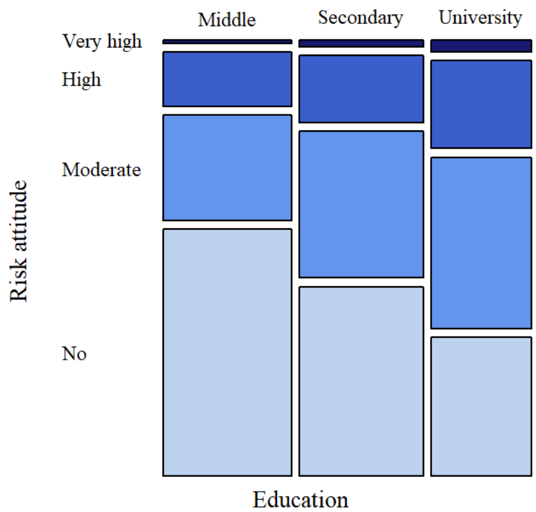


Source: Author's own work.

Financial literacy is assessed by resorting to the replies to two questions evaluating the basic knowledge of nominal and real interest rates. In particular, the binary variable $Literacy_1$ takes value 1 if the interviewed correctly answers to the question: 'Suppose you have 100 euros in a savings account and the interest rate is 2% per year. After 5 years, how much do you think you would have in the account if you left the money to grow?' (The answers are: (a) 'More than 110 euros', (b) 'Exactly 110 euros', (c) 'Less than 110 euros' and (d) 'Don't know'). $Literacy_2$ takes value 1 if the respondent correctly answers to the question: 'Imagine that the interest rate on your savings account is 1% per year and the inflation rate is 2% per year. After one year, with the money in the account, would you be able to buy (...)'. (The answers are: (a) 'More than today', (b) 'Exactly the same as today', (c) 'Less than today', (d) 'Don't know'). This way to measure financial competence is quite widespread in the literature dealing with financial literacy and the relationship with risk attitude (Lusardi & Mitchell, 2014; van Rooij et al., 2011b).

Figure 2 highlights that the distribution of risk attitude is likely to be affected by education. As the respondents' level of education rises, the share of those stating to belong to the no-risk class of investors clearly decreases. Conversely, the relative size of other classes increases. Figure 3 shows that financial literacy impacts the distribution of risk attitude as well. A rise in financial competence leads to a substantial increase in the share of those who choose risk class 2 (moderate risk), and a reduction for risk class 3 (high risk), with extreme classes 1 and 4 remaining substantially unchanged.

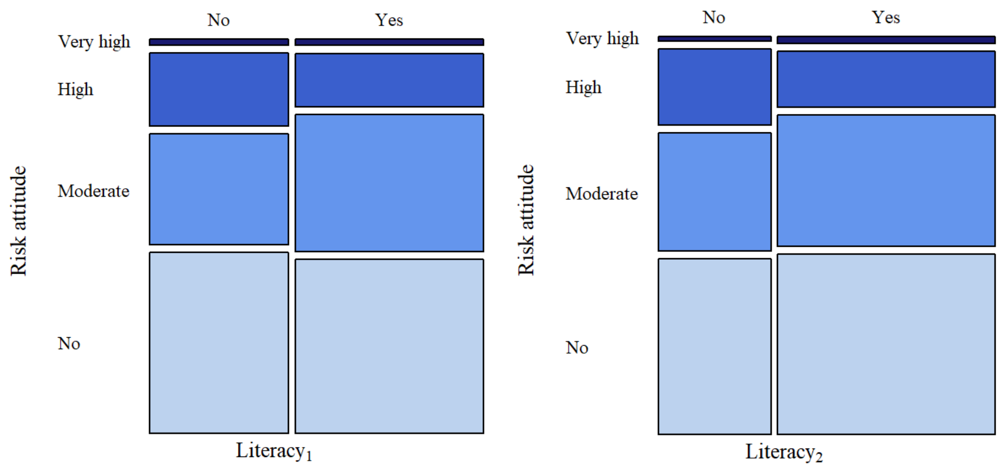
Figure 2: Distribution of risk attitude for education level



Source: Author's own work.

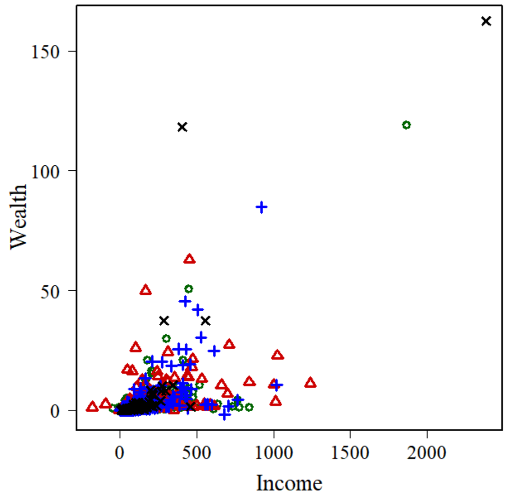
Concerning controls, the variables Age, Wealth and Income respectively denote age, family wealth and net individual disposable income of the respondent. The binary variable Hardship accounts for economic difficulties stated by the answer to the question: 'Is it difficult for your family to make ends meet?'. Urban is a binary variable taking unit value if the municipality where the household is located has at least 20.000 inhabitants. Finally, we introduce in the analysis the geographical variable CentreSouth accounting for the location of the household in one of the Central or Southern (rather than Northern) regions of Italy (The variable CentreSouth takes unit value for the regions Tuscany, Marche, Umbria, Latium, Abruzzo, Molise, Campania, Apulia, Basilicata, Calabria, Sicily and Sardinia, and zero value for Aosta Valley, Piedmont, Liguria, Lombardy, Trentino Alto Adige, Veneto and Friuli Venetia Giulia). Figure 4 displays the scatterplot of Income and Wealth, showing several outlying points.

Figure 3: Distribution of risk attitude for financial literacy



Source: Author's own work.

Figure 4: Wealth (million Euros) versus Income (thousand Euros)



*Responses are indicated by circles for no risk, triangles for moderate risk, + for high risk and X for very high risk

Source: Author's own work.

The statistical model

The self-reported risk attitude Y is an ordinal variable with m = 4 categories. Its relationship with the covariates is described through a cumulative model (McCullagh, 1980). Unlike most of the previous studies, which apply binary response models based on a dichotomization of Y, we choose to adopt the cumulative model to take into account the ordinal nature of the response, and avoid the loss of information produced by collapsing categories (Cohen, 1983).

The cumulative model assumes that the response derives from the categorization of an underlying (continuous) latent variable Y*, which in the current context is the actual risk attitude. The j-th category of Y is observed when Y* is in the interval (τ_j−1, τ_j) where

$$-\infty = \tau_0 < \tau_1 < \ldots < \tau_m = +\infty$$

are the thresholds on the support of Y*. Hence, for the i-th statistical unit, we have:

$$Y_i = j \Leftrightarrow \tau_{j-1} < Y_i^* \leq \tau_j,\ j = 1,\ldots,m \qquad (1)$$

Let $\mathbf{X}_i$ = (X_{i1}, … X_{ip}) be the corresponding vector of covariates. The latent variable Y_i* depends on $\mathbf{X}_i$ through the latent regression model:

$$\mathrm{Y}_i^* = \boldsymbol{X}_i \boldsymbol{b} + \mathrm{e}_i,\ \ \mathrm{i} = 1,\ldots,\mathrm{n} \qquad (2)$$

Where $\boldsymbol{\beta}$ = (β_1, ..., β_p)' is the vector of regression coefficients and is the error term.

The latent model for risk attitude is

$$\begin{aligned} Y_i^* &= \beta_1 Secondary_i + \beta_2 Tertiary_i + \beta_3 Literacy_{1i} + \\ &\beta_4 Literacy_{2i} + \beta_5 Age_i + \beta_6 Wealth_i + \beta_7 Income + \\ &\beta_8 (Income * Literacy_1)_i + \beta_9 (Income * Literacy_2)_i + \\ &\beta_{10} CentreSouth_i + \beta_{11} Hardship_i + \beta_{12} Urban_i + \varepsilon_i \end{aligned} \qquad (3)$$

From equations (1) and (2) the conditional cumulative probability of the response is

$$\mathrm{P}(\mathrm{Y}_i \leq \mathrm{j} | \boldsymbol{X}_i) = \mathrm{G}(\tau_j - \boldsymbol{X}_i \beta)$$

which yields

$$P(Y_i = j | X_i) = G(\tau_j - X_i\beta) - G(\tau_{j-1} - X_i\beta),\ j = 1,\ldots,m \quad (4)$$

A convenient choice for the distribution function of ε_i is the logistic distribution

$$G(u) = exp(u)/[1 + exp(u)]$$

which yields the cumulative logit model. It allows an easy assessment of the impact of covariates through the logit

$$\log\left\{\frac{\mathrm{P}(\mathrm{Y_i} \leq \mathrm{j}|\mathrm{X_i})}{\mathrm{P}(\mathrm{Y_i} > \mathrm{j}|\mathrm{X_i})}\right\} = \tau_j - \boldsymbol{X}_i \boldsymbol{\beta}$$

A positive β_k implies that the odds in favor of higher categories of risk attitude increase with the k-th covariate.

Estimation methods

Let

$$\boldsymbol{\tau} = (\tau_1, \ldots, \tau_{m-1})'$$

be the vector of thresholds, with the parameter

$$\boldsymbol{\theta} = (\boldsymbol{\tau}', \boldsymbol{\beta}')' \in \Omega(\theta)$$

Where Ω(**θ**) being an open subset of R^{p+m-1}. Consider a random sample of independent couples (Y_i, $\mathbf{X}_i$) for i = 1,…,n. The contribution of the i-th statistical unit to the log-likelihood function is

$$\ell(\boldsymbol{\theta}; Y_i, \boldsymbol{X}_i) = \sum_{j=1}^{m} \mathrm{II}(Y_i = j) lnP(Y_i = j | \boldsymbol{X}_i)$$

Where II(ω) is an indicator function, which takes value 1 if ω holds and 0 otherwise and ln(·) denotes the natural logarithm. It yields the log-likelihood function

$$\sum_{i=1}^{n} \ell(\theta; Y_i, X_i)$$

The score function is

$$S_n(\theta) = \sum_{i=1}^{n} S(\theta; Y_i, X_i)$$

Where

$$S(\boldsymbol{\theta}; Y_i, \boldsymbol{X_i}) = [S_{\tau_1}(\boldsymbol{\theta}; Y_i, \boldsymbol{X_i}), \ldots, S_{\tau_{m-1}}(\boldsymbol{\theta}; Y_i, \boldsymbol{X_i}), s_{\beta_1}(\boldsymbol{\theta}; Y_i, \boldsymbol{X_i}), \ldots, s_{\beta_p}(\boldsymbol{\theta}; Y_i, \boldsymbol{X_i})]' \textit{ and } S_{\theta_l}(\boldsymbol{\theta}; Y_i, \boldsymbol{X_i}) = \partial\theta_\ell(\boldsymbol{\theta}; \boldsymbol{Y_i}, \boldsymbol{X_i}) / \partial\theta_\ell$$

for l = 1, … m + p - 1. Denote by

$$g(u) = \exp(-u) / [1 + \exp(-u)]^2$$

the logistic density function. We have the thresholds

$$S_{n,\tau_s}(\boldsymbol{\theta}) = \sum_{i=1}^{n} S_{\tau_s}(\boldsymbol{\theta}; Y_i, \boldsymbol{X_i}) = \sum_{i=1}^{n} \frac{g(\tau_s - \boldsymbol{X_i\beta})[\mathrm{II}(Y_i = s) - \mathrm{II}(Y_i = s+1)]}{G(\tau_{s+1} - \boldsymbol{X_i\beta}) - G(\tau_s - \boldsymbol{X_i\beta})} \quad (5)$$

for s = 1, …, m-1, and regression coefficients

$$S_{n,\beta_k}(\boldsymbol{\theta}) = \sum_{i=1}^{n} S_{\beta k}(\boldsymbol{\theta}; Y_i, \boldsymbol{X_i}) = \sum_{i=1}^{n}\sum_{j=1}^{m} \mathrm{II}(Y_i = j) e_{ij}(\theta) X_{ik} \quad (6)$$

for k = 1, …, p, where $e_{ij} = (\theta)$ in (9) are the generalized residuals

$$e_{ij}(\boldsymbol{\theta}) = -\frac{g(\tau_j - \boldsymbol{X_i\beta}) - g(\tau_{j-1} - \boldsymbol{X_i\beta})}{G(\tau_j - \boldsymbol{X_i\beta}) - G(\tau_{j-1} - \boldsymbol{X_i\beta})}, \quad (7)$$

$$i = 1, \ldots, n; \; j = 1, \ldots, m$$

(see: Franses & Paap, 2004, and Iannario & Monti, 2023a, for their properties). The MLE is the solution

$$\widehat{\boldsymbol{\theta}} = (\widehat{\boldsymbol{\tau}}', \widehat{\boldsymbol{\beta}}')' \textit{ of } S_n(\widehat{\boldsymbol{\theta}}) = \mathbf{0}$$

Our purpose is to implement an estimation method yielding outlier stable estimators. In cumulative models anomalous data may be produced by either outlying covariates or anomalous responses, due to misclassifcation or incoherent respondents' behavior. To limit the impact of these data, robust estimators with a bounded influence function (Hampel et al., 1986) are needed. The influence function of MLEs is given by

$$IF(Y, \boldsymbol{X}; \boldsymbol{\theta}) = \Im(\boldsymbol{\theta})^{-1} S(\boldsymbol{\theta}; Y, \boldsymbol{X}), \textit{ where } \Im(\boldsymbol{\theta})$$

is the Fisher information matrix. Hence the influence function is bounded when all the elements of S(**θ**; **X**, Y) are bounded.

Iannario, Monti, Piccolo and Ronchetti (2017) and Scalera, Iannario and Monti (2021) show that in the cumulative logit model the score functions (5) of the thresholds are bounded and the generalized residuals vary within (−1,1). Unfortunately, expression (6) points out that an unlimited X_{ik} Iannario et al., (2017) and Scalera et al. (2021) show that in the cumulative logit model the score functions (8) of the thresholds are bounded and the generalized residuals vary within (−1,1). Unfortunately, expression (9) points out that an unlimited X_{ik} can make the score function of the regression coefficients unbounded. As a consequence, extreme design points can jeopardize the reliability of likelihood-based inference. This is a major concern in this context since both Wealth and Income are unlimited covariates, and indeed several extreme design points appear in Figure 4.

To achieve a bounded influence function estimation, in the case of outlying covariates, a weighted likelihood approach can be adopted (Croux et al., 2013; Iannario & Monti, 2023b). The individual contribution to the log-likelihood function is replaced by a weighted version

$$\ell_w(\boldsymbol{\theta}; Y_i, \boldsymbol{X_i}) = \ell(\boldsymbol{\theta}; Y_i, \boldsymbol{X_i}) w(\boldsymbol{X_i})$$

where w(**X**ᵢ) is a function downweighting extreme **X**ᵢ's. Equivalently the score function may be replaced by

$$\psi(\boldsymbol{\theta}; Y_i, \boldsymbol{X_i}) = S(\boldsymbol{\theta}; Y_i, \boldsymbol{X_i}) w(\boldsymbol{X_i})$$

This yields an M estimator $\widehat{\theta}_M$ of θ which is the implicit solution of

$$\sum_{i=1}^{n} \psi(\widehat{\boldsymbol{\theta}}_M; Y_i, \boldsymbol{X_i}) = 0 \quad (8)$$

To assess how much $\mathbf{X_i}$ is outlying, a Mahalanobis distance can be used. It is given by

$$\|\boldsymbol{X_i}\| = \left\{ \left(\boldsymbol{X_i} - \widetilde{\mu}_x\right)^T \widetilde{\sum}_X^{-1} (\boldsymbol{X_i} - \widetilde{\mu}_X) \right\}^{\frac{1}{2}} \quad (9)$$

Where $\widetilde{\boldsymbol{\mu}}_X$ and $\widetilde{\boldsymbol{\Sigma}}_X$ are robust estimators of the location and the covariance matrix of the covariates. In case the regressors are mixed, continuous (and potentially unlimited) and binary, the Mahalanobis distance is computed only on the continuous variables. Conditionally on the covariates, these estimators are Fisher consistent, i.e. under no contamination of the data we have

$$E\left\{\psi(\boldsymbol{\theta}; Y, X) \middle| \boldsymbol{X} = \boldsymbol{x}\right\} = \mathbf{0}$$

The weights need to be a decreasing function of $\|\mathbf{X_i}\|$. To this end, the Huber function (Hampel et al., 1986) can be used, leading to

$$w(\boldsymbol{X_i}) = w_i^X = min(1, c / \|\boldsymbol{X_i}\|)$$

When the tuning constant c increases, the M estimators approach the MLEs, whereas when c decreases, extreme design points are more severely downweighted. The square of the Mahalanobis distance $\|\mathbf{X_i}\|^2$ can be compared with the percentiles of a χ^2_p distribution, by considering as tuning constant

$$c = \sqrt{\chi^2_{p,\xi}}$$

Where $\chi_{p,\xi}$ is the ξ-th percentile of the χ_p^2 distribution. These weights have the advantage that they are computed only once at the beginning of the esti mation process and need not to be updated.

A more complex weight function takes into account the association between the covariates and the response. The denominator of the generalized residual in (7)

$$P(Y = j \,|\, \boldsymbol{X_i}), \textit{ hence } |e_{ij}(\boldsymbol{\theta})|$$

is large when the response has a low probability under the model, i.e. it is anomalous (Copas, 1988). Considering the product $|e_{ij}(\boldsymbol{\theta})|\|\mathbf{X}_i\|$ allows handling extreme design points differently according to whether the associated response is coherent with the model or not. It leads to the Huber weights (Iannario et al., 2017).

$$w_i^{eX} = w^{eX}(Y_i, \boldsymbol{X}_i \boldsymbol{\theta}) = min(1, c / \kappa(e_{ij}(\boldsymbol{\theta}), \boldsymbol{X}_i))$$

$$where\ \kappa(e_{ij}(\boldsymbol{\theta}), \boldsymbol{X}_i)) = \sum_{j=1}^{m} \text{II}(Y_i = j)\left|e_{ij}(\boldsymbol{\theta})\right|\left\|\boldsymbol{X}_i\right\|$$

These weights penalize more heavily outlying X_i whose response is anomalous. The M estimator, obtained with the weights w_i^{eX} is solution $\widehat{\boldsymbol{\theta}}_{\mathbf{M}}$ of (8) with

$$\psi(\boldsymbol{\theta}; Y_i, \boldsymbol{X}_i) = S(\boldsymbol{\theta}, Y_i, \boldsymbol{X}_i) w^{eX}(Y_i, \boldsymbol{X}_i \boldsymbol{\theta}) - \alpha(\boldsymbol{\theta})$$

where the term

$$\alpha(\boldsymbol{\theta}) = E\left[S\left(\boldsymbol{\theta}, Y_i, \boldsymbol{X}_i\right) w(Y, \boldsymbol{X}; \boldsymbol{\theta})\right]$$

is required to achieve Fisher consistency, i.e.

$$E\left[\psi\left(\boldsymbol{\theta}; Y, \boldsymbol{X}\right)\right] = \mathbf{0}$$

For both M estimators, the influence functions is

$$IF_M(y, \mathbf{x}; \boldsymbol{\theta}) = M_\psi(\boldsymbol{\theta})^{-1} \psi(\boldsymbol{\theta}; y, \mathbf{x})$$

for y = 1, …, m and $\mathbf{x} \in R^p$ where

$$M_\psi(\boldsymbol{\theta}) = -E\left\{\partial \psi(\boldsymbol{\theta}; Y, \boldsymbol{X}) / \partial \boldsymbol{\theta}\right\}$$

Under general regularity conditions (Huber, 1981), the M estimators are asymptotically normal, i.e.

$$n^{1/2}(\widehat{\boldsymbol{\theta}}_M - \boldsymbol{\theta}) \to N_{(m-1)(p+1)}(\mathbf{0}, \mathbf{V}_M(\boldsymbol{\theta}))$$

Where

$$\mathbf{V}_M(\boldsymbol{\theta}) = \boldsymbol{M}_\psi(\boldsymbol{\theta})^{-1} \mathbf{Q}_\psi(\boldsymbol{\theta}) \boldsymbol{M}_\psi^{-1}(\boldsymbol{\theta})$$

and

$$\mathbf{Q}_\psi(\boldsymbol{\theta}) = E\left[\psi(Y, \boldsymbol{X}; \boldsymbol{\theta}), \psi(Y, \boldsymbol{X}; \boldsymbol{\theta})^T\right]$$

Consequently, a test of hypothesis on a single parameter, i.e. H_0: $\theta_r = \theta_r^0$, can be carried out through a t-type statistic. We have

$$t_r = n^{1/2}(\widehat{\theta}_{M,r} - \theta_r^0) / (\widehat{V}_m^{rr})^{\frac{1}{2}} \to N(0,1)$$

Where $\widehat{V}_M^{RR}$ is the r-th element of the diagonal of the estimate $\widehat{\mathbf{V}}_{\mathbf{M}}$ of $\mathbf{V}_{\mathbf{M}}$.

Table 1: Maximum likelihood estimates

Variables	Coefficient	Standard Error	T-stat	P-value
Intercept 1|2 (τ1)	-0.611	0.157	-3.884	0.000
Intercept 2|3 (τ2)	1.091	0.158	6.912	0.000
Intercept 3|4 (τ3)	3.732	0.182	20.452	0.000
Secondary	0.332	0.064	5.172	0.000
Tertiary	0.552	0.075	7.367	0.000
$Literacy_1$	-0.239	0.056	-4.299	0.000
$Literacy_2$	-0.296	0.056	-5.247	0.000
Age	-0.014	0.002	-7.909	0.000
Wealth	0.008	0.002	3.557	0.000
Income	0.049	0.023	2.096	0.036
Income x $Literacy_1$	-0.023	0.021	-1.104	0.270
Income x $Literacy_2$	0.005	0.023	0.222	0.824
Centre South	0.485	0.051	9.474	0.000
Hardship	-0.551	0.058	-9.465	0.000
Urban	0.446	0.070	6.410	0.000

Source: Author's own work.

Empirical results

The cumulative logit model with the latent regression (3) has been estimated by MLEs and by the M estimators with weights W_i^X and W_i^{eX}. Since there are three continuous covariates, the tuning constant of the M estimator is

$$\sqrt{\chi^2_{3,0.7}} = 1.914$$

which produces a limited loss of efficiency in the case of pure data (Iannario et al., 2017; Iannario & Monti, 2023b). The Minimum Covariance Determinant estimators (Rousseeuw, 1984, 1985), which has a high breakdown point, have been applied for the estimators where μ_X and Σ_X in the Mahalanobis distance (9). As the ranges of Wealth and Income are extremely large, these covariates have been rescaled to aid a better convergence of the optimization algorithm. Let $\widetilde{X}_{ik}$ be the original value of the k-th covariate (Wealth or Income) corresponding to the i-th statistical unit. The rescaled values are given by

$$X_{ik} = (\widetilde{X}_{ik} - median(\widetilde{X}_k) / mad(\widetilde{X}_k))$$

Where median($\widetilde{X}_k$) and mad($\widetilde{X}_k$) are the median and the normalized median absolute deviation of $\widetilde{X}_k$. The ML estimates are in Table 1, while the M estimates with weights W_i^X and W_i^{eX} are in Tables 2 and 3 respectively.

Table 2: M estimates with weights w_i^X

Variables	Coefficient	Standard Error	T-stat	P-value
Intercept 1|2 (τ1)	-0.571	0.168	-3.392	0.001
Intercept 2|3 (τ2)	1.078	0.170	6.348	0.000
Intercept 3|4 (τ3)	4.024	0.210	19.111	0.000
Secondary	0.261	0.068	3.841	0.000
Tertiary	0.378	0.083	4.530	0.000
$Literacy_1$	-0.279	0.058	-4.767	0.000
$Literacy_2$	-0.308	0.059	-5.221	0.000
Age	-0.016	0.002	-8.021	0.000
Wealth	0.072	0.015	4.921	0.000
Income	0.255	0.051	4.975	0.000
Income x $Literacy_1$	-0.142	0.042	-3.346	0.001
Income x $Literacy_2$	-0.085	0.046	-1.855	0.064
Centre South	0.648	0.056	11.629	0.000
Hardship	-0.437	0.062	-7.053	0.000
Urban	0.475	0.071	6.645	0.000

Source: Author's own work.

Table 3: M estimates with weights w_i^{eX}

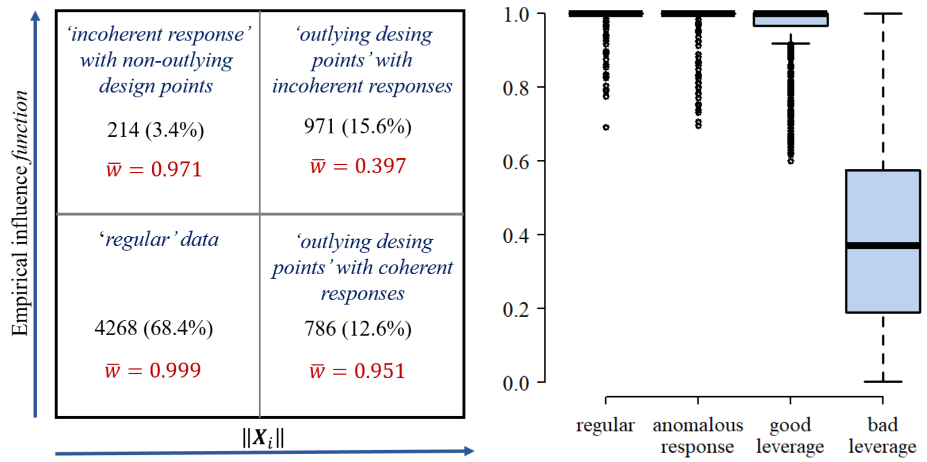
Variables	Coefficient	Standard Error	T-stat	P-value
Intercept 1|2 (τ1)	-0.603	0.163	-3.703	0.000
Intercept 2|3 (τ2)	1.094	0.165	6.642	0.000
Intercept 3|4 (τ3)	3.979	0.201	19.792	0.000
Secondary	0.281	0.067	4.173	0.000
Tertiary	0.411	0.081	5.104	0.000
$Literacy_1$	-0.264	0.058	-4.532	0.000
$Literacy_2$	-0.307	0.059	-5.214	0.000
Age	-0.015	0.002	-8.058	0.000
Wealth	0.068	0.010	6.896	0.000
Income	0.218	0.046	4.744	0.000
Income x $Literacy_1$	-0.109	0.037	-2.920	0.004
Income x $Literacy_2$	-0.078	0.038	-2.041	0.041
Centre South	0.599	0.054	11.137	0.000
Hardship	-0.488	0.061	-7.381	0.000
Urban	0.439	0.069	6.332	0.000

Source: Author's own work.

The main difference between ML and M estimates is given by the significance of the coefficients related to the interactions between financial literacy and income. The ML estimates are not significant. This may be due to the impact of the extreme design points which cause an implosion of the slope estimator toward zero (Croux et al., 2013). The M estimates instead highlight the relevance of the interactions, pointing out that financial knowledge mitigates the effect of an increase in income.

Pison and Van Aelst (2004) and Croux et al. (2013) suggest a procedure for detecting anomalous data that are influential for the MLEs, based on the Mahalanobis distance (9) and the empirical influence function. Details on this procedure and its application in the current context are given in the Appendix. This procedure allows us to classify the data in four groups: regular observations, anomalous responses, good leverage points characterized by an extreme design point with a response consistent with the model, and bad leverage points such that an extreme design point is associated with an anomalous response. The left panel of Figure 5 shows how many observations are identified as belonging to the four groups and, for each case, reports the average weight associated by M estimation with weights w_i^{eX}. Regular data have almost weight 1 on average, and no penalization is generally inflicted. The average weight of data whose response is identified as anomalous is $\overline{w} = 0.971$. Minimal down-weighting is applied to these data since they have a limited influence when the logistic link is adopted. Also, extreme design points associated with coherent responses are only mildly down-weighted, as the average weight is $\overline{w} = 0.951$. M estimation applies a severe down-weighting only to bad leverage points. These observations are very influential for the MLEs, but their influence on M estimators is strongly mitigated as $\overline{w}$ = 0.397. The boxplots of the weights for the four groups of data are displayed in the right panel of Figure 5.

Figure 5: Different kinds of observations and M weights w^{eX}



Source: Author's own work.

The empirical model obtained by the M estimators with weights W_i^{eX} is

$$\begin{aligned} logit = {} & \tau_j + \underset{(0.067)}{0.281}\,Secondary + \underset{(0.081)}{0.411}\,Tertiary - \\ & \underset{(0.058)}{0.264}\,Literacy_1 - \underset{(0.059)}{0.307}\,Literacy_2 - \underset{(0.002)}{0.015}\,Age + \\ & \underset{(0.010)}{0.068}\,Wealth + \underset{(0.046)}{0.218}\,Income - \underset{(0.037)}{0.109}\,Income \times \\ & Literacy_1 - \underset{(0.038)}{0.078}\,Income \times Literacy_2 + \\ & \underset{(0.054)}{0.599}\,CentreSouth - \underset{(0.061)}{0.488}\,Hardship + \underset{(0.069)}{0.439}\,Urban \end{aligned} \qquad (10)$$

All the explanatory variables considered in the analysis significantly affect risk attitude. In particular, the empirical model confirms the importance of education as a factor determining a stronger risk propensity, with a more intense impact of tertiary education. This outcome is obtained when both education and financial literacy are included among the covariates, but also if education is considered alone (i.e. without financial literacy).

The results also show that while overall education is a factor contributing to raising risk tolerance, financial literacy tends to reduce risk propensity. In addition, M estimators recognize statistical significance of the coefficients related to the income – financial literacy interaction terms. In this way, the moderating role of financial literacy is acknowledged for more affluent households, for whom the effect of financial competence on risk aversion is enhanced.

Concerning control variables, the coefficients show signs consistent with the previous studies investigating the impact on risk aversion of personal, social and context factors. In particular, as previously highlighted by many contributions (e. g. Bellante & Green, 2004; Yao et al., 2011), risk propensity turns out to be negatively related to age. Older people are more unwilling to take risks for psychological reasons. Furthermore, age is positively associated to perspectives of worse health and lower future incomes, which in turn inhibit risk propensity. Confirming a consolidated result of the literature, wealth and income (Dohmen et al., 2011; Guiso & Paiella, 2008; Hallahane al., 2004) on one side, and economic uncertainty (Hochguertel, 2003) on the other, are significantly related to risk propensity, respectively with positive and negative signs. We also find that urbanization is relevant to risk attitude, as urbanized population tends to be less risk-averse, a result similar to those documented by Outreville (2015) as well as Shi and Yan (2018). Finally, an interesting finding is that, despite the lower per capita income and wealth, residents in the Centre and the South show a higher propensity for risk. A lower average age and a higher share of self-employed workers may help explain the observed lower risk aversion in the Centre-South.

Figure 6 shows the estimated probability distribution of risk attitude, based on (10), when income and wealth vary. Five profiles are considered by setting income and wealth at their η-th percentile, where η = 5, 25, 50, 75, 90. Their values are reported in Table 4. The left panel considers the case of financial illiteracy ($Literacy_1 = Literacy_2 = 0$) while in the right panel financial literacy is assumed ($Literacy_1 = Literacy_2 = 1$). The individual is assumed to have a secondary school degree, live in an urbanized area of Northern Italy, with his/her family having no economic hardship. Age is set at its median value (62 years).

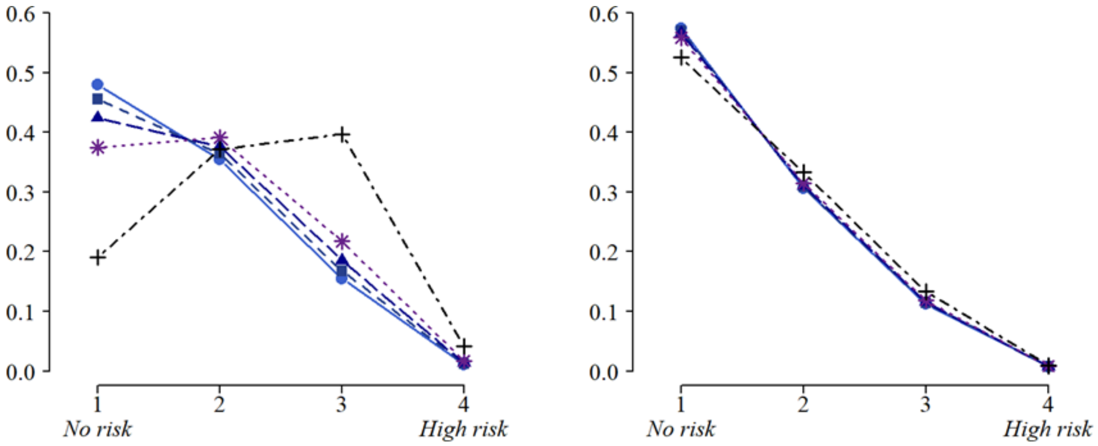
In the case of financial illiteracy, comparatively higher probabilities are assigned to the moderate - or high-risk classes, and the moderate-risk or even high-risk class becomes the modal one for high-income households. More financial competence brings about a more cautious attitude (the probability of respondents being in no - or moderate- risk classes exceeds 0.85). In addition, differences in the distribution of risk attitude across income classes tend to vanish. The same pattern emerges when the distribution is studied for different percentiles of age and educational levels, as well as across Northern and Centre-Southern households, and residents in urban and rural areas (Evidence not reported here is available upon request).

Table 4: Percentiles of Income and Wealth (thousands Euros)

ξ	5	25	50	75	95
Income	10.2	21.9	37.2	61.8	173.5
Wealth	1.0	92.0	221.5	501.1	221.4

Source: Author's own work.

Figure 6: Distribution of risk attitude at selected percentiles of Income and Wealth



*5th (solid line with circles), 25th (short-dashed line with squares), 50th (long-dashed line with triangles), 75th (dotted line with stars) and 95th percentile (dot-dashed line with plus). Lack of financial literacy in the left panel and financial literacy in the right panel (other covariates: median Age, North, Secondary School, No hardship, Urbanization)

Source: Author's own work.

Final remarks

To assess the role of education and financial literacy on self-reported risk attitude, a cumulative logit model is applied. Estimation is performed through MLEs and M estimators with a bounded influence function. The latter effectively limits the impact of anomalous data on the fitted model, leading to a proper assessment of the effect of covariates.

The empirical model shows that overall education contributes to raise risk tolerance, while financial literacy tends to reduce risk propensity.

Appendix

This appendix illustrates the procedures leading to the classification of the data displayed in Figure 5.

The information matrix is given by

$$\Im(\boldsymbol{\theta}) = E_x\left[E_Y\left[S(\boldsymbol{\theta},Y,\boldsymbol{X})S(\boldsymbol{\theta},Y,\boldsymbol{X})^T\,|\boldsymbol{X}\right]\right]$$

where the conditional expectation of the product of the score function given X = x is

$$E_Y\left[S(\boldsymbol{\theta},Y,\boldsymbol{X})S(\theta,Y,\boldsymbol{X})^T\,|\boldsymbol{X}=\boldsymbol{x}\right]=$$
$$\sum\nolimits_{j=1}^{m} S(\boldsymbol{\theta};j,\boldsymbol{x})S(\boldsymbol{\theta};j,\boldsymbol{x})^T P(Y=j|\boldsymbol{x})$$

In the observed information matrix the parameter **θ** is replaced by the robust estimator $\widehat{\boldsymbol{\theta}}_{\mathbf{M}}$, which is unaffected by outliers. By following Croux et al. (2013), the distribution of covariates is estimated by the empirical distribution based on the observations which are not detected as outliers in the space of the continuous explanatory variables, i.e. for which

$$\|\boldsymbol{X}_i\| \le \sqrt{\chi^2_{pc,\xi}} = d$$

where p_c is the number of continuous variables and ζ = 0.975. Hence the observed information matrix is computed as follows

$$\widehat{\Im}(\widehat{\boldsymbol{\theta}}_M) = \sum\nolimits_{\|X_i\|\le d} \sum\nolimits_{j=1}^{m} S(\widehat{\boldsymbol{\theta}}_M; j,\boldsymbol{X}_i)S(\widehat{\boldsymbol{\theta}}_M; j,\boldsymbol{X}_i)^T$$
$$\widehat{P}_{\boldsymbol{\theta}_M}(Y=j|\boldsymbol{X}_i)$$

where

$$\widehat{P}_{\widehat{\boldsymbol{\theta}}_M}(Y=j|\boldsymbol{X}_i)$$

is the estimated conditional probability obtained by replacing θ by $\widehat{\boldsymbol{\theta}}_{\mathbf{M}}$ in equation (4).

The empirical influence function for the MLE corresponding to (Y_i, $\mathbf{X}_i$) is

$$EIF(Y_i,\boldsymbol{X}_i) = \widehat{\Im}(\widehat{\theta}_M)^{-1}S(\widehat{\theta}_M;Y_i,X_i)$$

The overall influence measure of a single point (Y_i, $\mathbf{X}_i$) on the regression coefficients is computed as the Euclidean norm of the sub-vector of the influence function EIFβ (y, **x**) related to β standardized by the number of parameters (Pison & Van Aelst, 2004; Croux et al., 2013)

$$\overline{EIF}_{\beta,i} = \frac{EIF_{\beta}(Y_i, X_i)^T EIF_{\beta}(Y_i, X_i)}{\sqrt{p}}$$

To detect influential points, it is necessary to identify a cutoff for EIF_i. The Monte Carlo procedure by Pison and Van Aelst (2004) is adapted to the present context through resampling. B = 1000 datasets of the same size n are generated by the model fitted by M estimators. The covariates are sampled from the estimated distribution of the $\mathbf{X}_i$. Let (X_i*, Y_i*) for i=1, 2, …, n be a bootstrap sample and let $\widehat{\boldsymbol{\theta}}_M^*$ be the M estimator. For each dataset the empirical influence function of (Y_i*, $\mathbf{X}_i$*) is replicated as

$$EIF(Y_i^*, X_i^*) = \widehat{\Im}(\widehat{\theta}_M^*)^{-1} S(\widehat{\theta}_M^*; Y_i^*, X_i^*)$$

Where

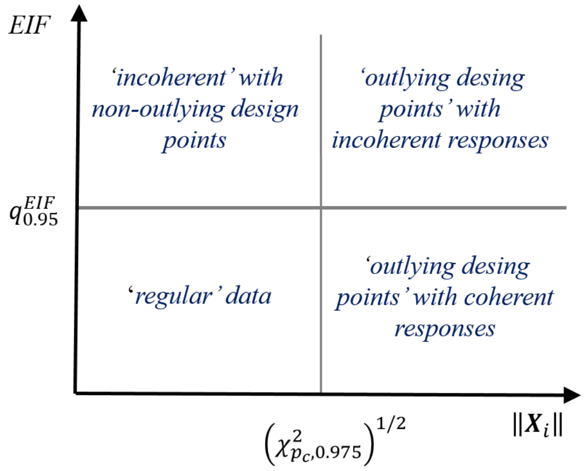
$$\widehat{\Im}(\widehat{\theta}_M^*) = \sum_{i=1}^{n} \sum_{j=1}^{m} S(\widehat{\theta}_M^*; j, X_i^*) S(\widehat{\theta}_M^*; j, X_i^*)^T$$

$$\widehat{P}_{\theta_M^*}(Y = j \mid X_i^*)$$

The information matrix is computed by summing over all the $\mathbf{X}_i$*, since resampling is performed from the empirical distribution of $\mathbf{X}_i$, after removing outlying points. The cutoff $q_{0.95}^{EIF}$ is determined as the 95th percentiles of all the overall empirical influence functions obtained for the observations of all the B datasets.

This procedure allows to split the data in four groups (Pison & Van Aelst, 2004) as illustrated in Figure 7. Observations such that $\|\mathbf{X}_i\| \le d$ and $EIF_i \le q_{0.95}^{EIF}$ are regular observations. In case of anomalous response we have $\|\mathbf{X}_i\| \le d$ and $EIF_i > q_{0.95}^{EIF}$. For good leverage points we have $\|\mathbf{X}_i\| > d$ and $EIF_i < q_{0.95}^{EIF}$. Finally, bad leverage points are such that $\|\mathbf{X}_i\| > d$ and $EIF_i > q_{0.95}^{EIF}$.

Figure 7: Observations according to their Mahalanobis distance and their empirical influence function



Source: Author's own work.

Acknowledgments

Maria Iannario gratefully acknowledges the support of the Marie Skłodowska-Curie Actions under the European Union's Horizon Europe research and innovation program for the Industrial Doctoral Network on Digital Finance (acronym: DIGITAL, project no. 101119635), and the Cost Action CA19130 "Fintech and Artificial Intelligence in Finance - Towards a transparent financial industry" (FinAI), supported by COST (European Cooperation in Science and Technology).

References

Bannier, C.E. & Neubert, M. (2016). Gender Differences in Financial Risk Taking: The Role of Financial Literacy and Risk Tolerance. Economic Letters, 145, 130-135.

Barsky, R.B., Juster, T.F., Kimball, M.S. & Shapiro, M.D. (1997). Preference Parameters and Behavioral Heterogeneity: An Experimental Approach in the Health and Retirement Study. Quarterly Journal of Economics, 112(2), 537-579.

Bayar, Y., Sezgin, H.F., Ozturk, O.F. & Sasmaz, M.U. (2020). Financial Literacy and Financial Risk Tolerance of Individual Investors: Multinomial Logistic Regression Approach. SAGE Open Journals, 10(3), 1-11.

Bayer, P.J., Bernheim, B.D. & Scholz, J.K. (2009). The Effects of Financial Education in the Workplace: Evidence from a Survey of Employers. Economic Inquiry, 47(4), 605-624.

Bellante, D. & Green, C.A. (2004). Relative Risk Aversion among the Elderly. Review of Financial Economics, 13(3), 269-281.

Browne, M.J., Chung, J. & Frees, E.W. (2000). International Property-Liability Insurance Con sumption. Journal of Risk and Insurance, 67(1), 73-90.

Cohen, J. (1983). The Cost of Dichotomization. Applied Psychological Measurement, 7, 249-53.

Chong, A. & Martinez, J.J. (2019). Does education increase risk aversion? Evidence Using Artefactual Experiments in Peru. International Center for Public Policy. Working Paper 19-17. Georgia State University, Georgia.

Cohn, R.A., Lewellen, W.G., Lease, R.C. & Schlarbaum, G.G. (1975). Individual Investor Risk Aversion and Investment Portfolio Composition. The Journal of Finance, 30, 605-620.

Copas, J.B. (1988). Binary Regression Models for Contaminated Data. Journal of the Royal Statistical Society: Series B, 50, 225-265.

Courbage, C., Montoliu-Montes, G. & Rey, B. (2018), How Vulnerable Is Risk Aversion to Wealth, Health and Other Risks? An Empirical Analysis for Europe, https://shs.hal.science/halshs-01935846/file/1827.pdf (Accessed: 08.02.2024).

Croux, C., Haesbroeck, G. & Ruwet, C. (2013). Robust Estimation for Ordinal Regression. Journal of Statistical Planning and Inference, 143, 1486-1499.

De Paola, M. (2012). The Determinants of Risk Aversion: The Role of Intergenerational Transmission. German Economic Review, 14(2), 214-234.

Dimmock, S.G., Kouwenberg, R., Mitchell, O.S. & Peijnenburg, K. (2016). Ambiguity Aversion and Household Portfolio Choice Puzzles: Empirical Evidence. Journal of Financial Economics, 119, 559-577.

Doepke, M. & Zilibotti, F. (2005). Social Class and the Spirit of Capitalism. Journal of the European Economic Association, 3, 516-524.

Dohmen, T., Falk, A., Huffman, D., Sunde, U., Schupp, J. & Wagner, G.G. (2011). Individual Risk Attitudes: Measurement, Determinants and Behavioral Consequences. Journal of the European Economic Association, 9(3), 522-550.

Fisher, P.J. & Yao, R. (2017). Gender Differences in Financial Risk Tolerance. Journal of Economic Psychology, 61, 191-202.

Franses, P.H. & Paap, R. (2004). Quantitative Models in Marketing Research, Cambridge University Press, Cambridge.

Friend, I. & Blume, M.E. (1975). The Demand for Risky Assets. The American Economic Review, 65, 900-922.

Guiso, L. & Paiella, M. (2008). Risk Aversion, Wealth, and Background Risk. Journal of the European Economic Association, 6, 1109-1150.

Halek, M. & Eisenhauer, J.G. (2001). Demography of Risk Aversion. Journal of Risk and Insurance, 68(1), 1–24.

Hallahan, T.A., Faff, R.W. & McKenzie, M.D. (2004). An Empirical Investigation of Personal Financial Risk Tolerance. Financial Services Review, 13, 57-78.

Hampel, F.R., Ronchetti, E.M., Rousseeuw, P.J. & Stahel, W.A. (1986). Robust Statistics: The Approach Based on Influence Functions. John Wiley & Sons, New York.

Harrison, G.W., Lau, M.I. & Rutstrom, E.E. (2007). Estimating Risk Attitudes in Denmark: A Field Experiment. Scandinavian Journal of Economics, 109(2), 341-368.

Hermansson, C. & Johnson, S. (2021). The Impact of Financial Literacy and Financial Interest on Risk Tolerance. Journal of Behavioral and Experimental Finance, 29, 1-12.

Hersch, J. (1996). Smoking, Seat Belts, and Other Risky Consumer Decisions: Differences by Gender and Race. Managerial and Decision Economics, 17(5), 471-481.

Hochguertel, S. (2003). Precautionary Motives and Portfolio Decisions. Journal of Applied Econometrics, 18, 61-77.

Huber, P.J. (1981). Robust Statistics. John Wiley & Sons, New York.

Hwang, T. & Gao, S. (2003). The Determinants of the Demand for Life Insurance in an Emerging Economy. The Case of China. Managerial Finance, 29(5/6), 82-96.

Iannario, M., Monti, A.C., Piccolo, D. & Ronchetti, E. (2017). Robust Inference for Ordinal Response Models. Electronic Journal of Statistics, 11, 3407-3445.

Iannario, M. & Monti, A.C. (2023a). Generalized Residuals and Outlier Detection for Ordinal Data with Challenging Data Structures. Statistical Methods & Applications, 32, 1197-1216.

Iannario, M. & Monti, A.C. (2023b). Robust Logistic Regression for Ordered and Unordered Responses. Econometrics and Statistics.

Jianakoplos, N.A. & Bernasek, A. (1998). Financial Risk Taking by Age and Birth Cohort. Southern Economic Journal, 72(4), 981-1001.

Jung, S. (2015). Does Education Affect Risk Aversion? Evidence from the British Education Reform. Applied Economics, 47(28), 2924-2938.

Kapteyn, A. & Teppa, F. (2011). Subjective Measures of Risk Aversion, Fixed Costs, and Portfolio Choice. Journal of Economic Psychology, 32(4), 564-580.

Li, J., Li, Q. & Wei, X. (2020). Financial Literacy, Household Portfolio Choice and Investment Return. Pacific-Basin Finance Journal, 62, 101370.

Lin, F.T. (2009). Does the Risk Aversion Vary with Different Background Risk of Households? International Research Journal of Finance and Economics, 34, 69-82.

Lusardi, A. & Mitchell, O.S. (2014). The Economic Importance of Financial Literacy: Theory and Evidence. Journal of Economic Literature, 52(1), 5-44.

Lusardi, A. & Mitchell, O.S. (2023). The Importance of Financial Literacy: Opening a New Field. Journal of Economic Perspectives, 37(4), 137-154.

McCullagh, P. (1980). Regression Models for Ordinal Data (with Discussion). Journal of the Royal Statistical Society, Series B, 42, 109-142.

Outreville, J.F. (2015). The Relationship between Relative Risk Aversion and the Level of Education: A Survey and Implications for the Demand for Life Insurance. Journal of Economic Surveys, 29(1), 97-111.

Pison, G. & Van Aelst, S. (2004). Diagnostic Plots for Robust Multivariate Methods. Journal of Computational and Graphical Statistics, 13, 310-329.

Rousseeuw, P.J. (1984). Least Median of Squares regression. Journal of the American Statistical Association, 79, 871-880.

Rousseeuw, P.J. (1985). Multivariate Estimation with High Breakdown Point. In: W. Grossmann, G. Pflug, I. Vincze, W. Wertz (Eds.), Mathematical Statistics and Applications (pp. 283-297). Dordrecht, Reidel.

Scalera, V., Iannario, M. & Monti, A.C. (2021). Robust link functions. Statistics, 55, 963-977.

Shi, X. & Yan, Z. (2018). Urbanization and Risk Preference in China: A Decomposition of Self-Selection and Assimilation Effects. China Economic Review, 49, 210-228.

Stoian, A., Vintila, N., Iorgulescu, F., Cepoi, C.O. & Dina Manolache, A. (2021). How Risk Aversion and Financial Literacy Shape Young Adults' Investment Preferences. MPRA Paper 109755.

Sutter, M., Weyland, M., Untertrifaller, A. & Froitzheim, M. (2020). Financial Literacy, Risk and Time Preferences. Results from a Randomized Educational Intervention. ECONtribute Discussion Paper 19. University of Bonn and University of Cologne, Germany.

Van Rooij, M.C.J., Lusardi, A. & Alessie R.J.M. (2011a). Financial Literacy and Retirement Planning in the Netherlands. Journal of Economic Psychology, 32, 593-608.

Van Rooij, M.C.J., Lusardi, A. & Alessie, R.J.M. (2011b). Financial Literacy and Stock Market Participation. Journal of Financial Economics, 101(2), 449-472.

Yao, R., Sharpe, D.L. & Wang, F. (2011). Decomposing the Age Effect on Risk Tolerance. Journal of Socio-Economics, 40, 879-887.