

RESEARCH

Open Access



The challenge of assessing financial literacy: alternative data analysis methods within the Italian context

Paola Bongini¹, Paola Iannello^{2*} , Emanuela E. Rinaldi¹, Mariangela Zenga³ and Alessandro Antonietti²

*Correspondence:

paola.iannello@unicatt.it

² Department of Psychology,
Università Cattolica del Sacro
Cuore, Largo Gemelli, 1,
20123 Milan, Italy

Full list of author information
is available at the end of the
article

Abstract

Background: Assessing individuals' financial literacy levels is currently widely recognized as being necessary to design effective financial education programs and also to evaluate their actual impact. To address the lack of a consensus regarding an appropriate instrument to measure financial literacy, the OECD and its International Network on Financial Education (INFE) developed a core questionnaire in 2011, to be administered across a wide range of countries. Italy participated in the study with a survey promoted by the financial consortium ABI-PattiChiari. A tailored version of the OECD/INFE questionnaire was used in the survey, with three indicators of financial literacy taken from the OECD survey (financial behavior index, financial attitude index, financial knowledge index) and two new indicators (financial familiarity index and financial planning).

Purpose: The present paper focuses on data analysis methods used to evaluate financial literacy among the Italian adult population. It reviews data analysis approaches used to evaluate financial literacy and proposes a new method to gauge this latent construct in order to obtain a valid and reliable index that is able to capture educational needs in a manner that is as accurate and targeted as possible.

Methods: The sample used for the survey consisted of 1247 Italian residents of at least 18 years of age who were reached via CATI. The sample was obtained by appropriate stratification across several dimensions (gender, age, geographical area, and municipality size). We propose alternative data analysis methods to treat the survey data: item response theory (IRT) and classification and regression tree analysis.

Results: The analysis highlighted the crucial role that data analysis methods play in assessing financial literacy. Comparing the results for classical test theory and IRT, this paper suggests that financial literacy research should be open to alternative and multiple approaches to obtain reliable measures of financial literacy that are able to capture the educational needs of different population groups and can help to design effective financial education programs.

Keywords: Financial literacy index, IRT analysis, CART analysis

Introduction

Buying a home or ensuring an income for one's retirement are just a couple of examples of the many situations that individuals will face during their lives and which require basic financial knowledge to make sound decisions. Currently, the ability to make informed, aware, and efficient financial decisions seems to be particularly important. Some recent trends converge to demonstrate a real need to promote and improve individuals' financial literacy, especially in some countries, such as Italy (Coppola et al. 2017). As pointed out by analyses conducted in Italy and elsewhere (see The European House-Ambrosetti, Consorzio PattiChiari 2007; Grifoni and Messy 2012; Lusardi and Mitchell 2014), compared to past generations, people live longer (and live longer in retirement), which clearly suggests that there is a need to effectively manage money to achieve life-long financial security. In this regard, another very recent trend is the greater personal responsibility that individuals have over pension planning, and the anxiety that they associate with it (Nicolini 2017). In Italy, the willingness to handle financial information is less pronounced for women, seniors and the unemployed (sometimes considered as "more vulnerable groups" in terms of poverty risk). Additionally, a low level of financial knowledge, financial anxiety and a lack of interest in financial issues show a negative correlation (Linciano et al. 2017). Taken together, these factors create challenges that might be more easily faced with the aid of financial education.

It is currently acknowledged that the design of effective educational interventions needs to be preceded by a thorough assessment of the initial or baseline level of financial literacy. This initial step is crucial for two reasons. First, a sound assessment approach helps in correctly identifying the educational gaps or biases to be addressed by the education programs. Second, this initial assessment represents an indispensable prerequisite for evaluating the success and impact of the defined interventions. Thus, assessing financial literacy is a challenge that needs to be undertaken by any organization, either public or private, that wants to engage in financial education at any level.

To date, the issue of assessing financial literacy has concentrated on the conceptual definition of the latent variable called "financial literacy". Huston (2010) and Remund (2010) provide a thorough literature review that helps frame the issue of defining the concept of what financial literacy is or should be. Indeed, financial literacy has been variably defined as specifically referring to a form of knowledge (e.g., Hilgert et al. 2003), the ability to apply that knowledge (e.g., Mandell 2007), and good financial behavior (e.g., Moore 2003). The methods used to measure financial literacy vary quite substantially according to the different conceptual definitions adopted. In fact, without a consensual definition, financial literacy has been measured dissimilarly across studies. The construct either focuses on a few financial issues or covers a wide variety of financial topics, including debt, insurance, spending, budgeting, inflation, investments, and saving for retirement. Analogously, the number of questions used to assess financial literacy levels also varies widely, ranging from 3 to 45 items. Across studies, both performance tests (usually multiple-choice questionnaires) and self-report methods have been employed to measure financial literacy. Performance tests are mainly knowledge-based (e.g., Mandell 2007), while self-reports tend to assess perceived knowledge. More recently, tests have been designed to gauge both objective knowledge and perceived knowledge. In general, considerable progress has been achieved in the design of surveys aimed at identifying

individual levels of financial literacy through the effort made by the OECD and its International Network on Financial Education (INFE) to develop and promote a common questionnaire based on the experience of a large number of previous rigorous national and international surveys. The OECD/INFE questionnaire and the underlying approach were described in OECD INFE (2011) and discussed in great detail by Kempson (2009). The questionnaire has been used in many countries (Atkinson and Messy 2012) to collect comparable data, including Italy (ABI-PattiChiari 2014).

In contrast, until very recently, the process of data analysis (i.e., of analyzing the information obtained through the questionnaires) has been less explored. Both bivariate and multivariate techniques are usually applied. In general, the responses to the proposed questions are simply summed to generate a score of financial literacy, which typically ranges between zero and the maximum number of correct answers. More recent studies have applied factor analysis (van Rooij et al. 2011). It is widely acknowledged, however, that more work is needed to develop rigorous psychometric analysis (Knoll and Houtts 2012). Leveraging the Italian experience in assessing financial literacy at the national level, this paper critically reviews data analysis approaches used to evaluate financial literacy, and proposes a new method to gauge this latent construct in order to obtain a valid and reliable index that is able to capture educational needs in a manner that is as accurate and targeted as possible.

The remainder of this paper is organized as follows. The second section provides an overview of the OCSE-INFE international survey on financial literacy and the main results obtained when run in Italy in 2013. Section three describes the sample used in the current study and outlines our approach to data analysis. Section four presents the empirical results. The final section summarizes the main results and draws some conclusions.

Background: the OECD–INFE international survey on financial literacy

In 2011 the OECD promoted a financial literacy survey within the framework of the INFE (OECD INFE 2011). The aim of the project was to collect information on the level of financial literacy among member countries, fully consistent with OECD recommendations to guarantee the cross-country comparability of financial literacy indicators. The findings of the first pilot study based on this approach are illustrated in Atkinson and Messy (2012).

Italy joined the survey in 2013 by means of a public–private partnership led by “PattiChiari,” a consortium of Italian banks committed to promoting market transparency and financial education. The questionnaire used to assess the respondents’ level of financial literacy closely followed the OECD INFE guidelines to measure financial literacy across countries. It included both a core questionnaire and a set of supplementary questions aimed at investigating issues such as the ability to properly access and use financial information and to plan for retirement, which were considered to be of interest for a better description of the Italian population’s level of financial literacy.

Following the OECD approach detailed in Atkinson and Messy (2012), the pieces of information provided in the core sections of the questionnaire were used to define three indicators of financial literacy: a financial behavior index (FBI), financial attitude index (FAI), and financial knowledge index (FKI). Furthermore, by exploiting the supplementary questions of the survey, two further indicators were created: a financial familiarity index (FFI; measuring the knowledge and usage of financial products

and services), and financial planning index (FPI; focusing on the respondents' ability to plan for retirement). The exact definitions and statistical descriptions of all indicators are reported in ABI-PattiChiari (2014) and Baglioni et al. (2018).

The FBI was obtained as an additive indicator, ranging from zero to nine, based on the answers to questions focusing on the financial decisions of the respondent, and assigning a score to each answer that increased with the quality of each decision (i.e., "savvy" financial behavior). The questions covered the consistency of the respondent's purchases with respect to budget constraints, the ability to meet payment deadlines and to maintain an adequate financial budget, the quality of savings and the choices of financial products, and the ability to commit to long-term financial planning.

The FAI provided a measure of the respondents' propensity to save. The index was built by assessing the individual's attitude toward saving for the future, as well as the perception of the tradeoff between current and future spending. The index was obtained by means of categorical scores (between 1 and 5) with higher values indicating a higher propensity to save.

Following Lusardi and Mitchell (2011), the FKI was based on the number of correct answers given by the respondent to questions addressing simple financial concepts such as the role of inflation, the ability to compute simple and compound interest rates, the relationship between risk and return, and the notion of portfolio diversification.

The two supplementary indicators provided a closer look into the respondent's familiarity with financial products and the ability to plan for retirement. The FFI was based on the respondent's knowledge and usage of fifteen financial instruments, ranging from bank accounts and credit cards to mutual funds, stocks and shares, and insurance products.

The FPI aimed to establish the respondents' awareness of the necessity to plan savings in advance in order to smooth consumption over the entire life cycle. The index was based on three questions assessing the existence of a financial budget at the household level, familiarity with supplementary pension funds, and familiarity with other forms of long-term savings to support retirement income.

To obtain a comprehensive measurement of financial literacy, the indexes described above were then aggregated, building on the approach detailed by Atkinson and Messy (2012). Indeed, the authors highlight that "financial literacy is a combination of knowledge, attitude and behavior, and so it makes sense to explore these three components in combination [...] by adding the scores together" (Atkinson and Messy 2012, p. 39). The financial literacy index was therefore built as a simple average of the three indexes describing an individual's financial behavior, financial attitude and financial knowledge. Along with this first financial literacy index, a second and more comprehensive financial literacy index was computed. This included all five elementary indexes depicted above; i.e., the three indexes suggested in the OECD guidelines plus the two indexes obtained from the supplementary questions.

Once the elementary and comprehensive indexes had been computed, they were subsequently used to analyze their relationship with the usual set of sociodemographic and economic characteristics of the respondents to obtain a description of the determinants of the level of financial literacy of the Italian population.

Two main methods were applied: ordered probit and ordinary least square (OLS) regressions¹ and classification and regression tree (CART) analysis.² The former are traditional forms of analysis that estimate the relationships between a dependent variable (which could be a categorical and ordinal variable or continuous variable) and a set of independent variables. The latter is a non-parametric regression and classification method originally introduced by Breiman et al. (1984). It allows the simultaneous identification of significant covariates impacting on the dependent variable of interest (in our case, the financial literacy indexes) and significant clusters (in our case, of individuals) that exhibit relevant differences with respect to the dependent variable, and homogeneous characteristics with respect to the explanatory variables considered.

In other words, using probit or OLS regressions, the researcher obtains a causal relationship (significance and relevance of impact) between the level of financial literacy and the sociodemographic explanatory variables under investigation. With CART analysis, the researcher is also able to split the sample into relevant and homogeneous clusters that exhibit differences in their sociodemographic characteristics with respect to their (similar) level of financial literacy.

These two diverse approaches to data analysis produce different results concerning the main determinants of the aggregate level of financial literacy as well as its elementary factors (knowledge, behavior, attitude).

Tables 5 and 6 and Figs. 3, 4 and 5 in Appendix 1 present these differences. The tables include the results of ordered probit and OLS regressions applied to each elementary index and to the aggregate indicators of financial literacy.³ Figures 3, 4, 5 illustrate the results of applying CART analysis. The differences are immediately apparent in two ways. First, the influencing covariates are not necessarily the same. Second, with CART analysis it is possible to identify different clusters of individuals with respect to the same variable that regression analysis highlighted as being relevant in influencing their level of financial literacy.

For instance, with respect to the FFI (Fig. 3), the CART analysis resulted in ten different clusters of individuals, initially identified with respect to their participation in the labor force. In this case, what was then relevant in explaining their familiarity with financial products was their level of income, followed by the area of residence for low-income individuals. On the other hand, for inactive individuals (in search of a job but also retirees and students), the second most important explanatory variable was the area of residence, followed by education for those living in southern Italy; however, marital status was more important for individuals living in northern Italy. On the other hand, the probit regression applied to the FFI resulted in a larger number of influencing factors beyond those highlighted by the CART analysis; i.e., gender, age, and direct involvement in the financial decisions of the household. In addition, regression analysis suggested to policy makers that all individuals sharing the same characteristics (for instance, living in southern Italy) are potentially identical targets for the same education program, showing the same deficit in financial literacy. In

¹ See Chapter 7, ABI-PattiChiari (2014) and Baglioni et al. (2018).

² Chapter 3, ABI-PattiChiari (2014).

³ See Chapter 7, ABI-PattiChiari (2014) and Baglioni et al. (2018) for the whole set of specifications estimated. For the sake of brevity, we report the two most comprehensive models, including all possible covariates and their meaningful interactions. In particular, these specifications control for whether the respondent plays an active role in making the household's financial decisions and check for the robustness of the gender effect (including a dummy for the marital status of the respondent and its interaction with gender).

contrast, CART analysis showed that there are at least three very different clusters among individuals residing in southern Italy (those active in the job market, those not working with low education, and those inactive with a higher level of education), suggesting that individuals not working and with a low level of education comprise the target group most in need of financial education programs.

Similarly, considering the financial knowledge of the respondents (Fig. 4), CART analysis identified gender as the first discriminant factor in the sample of respondents, followed by education and income. As before, the regression analysis identified, with no scale of priority, a larger number of significant covariates, including age and having an active role in financial decision making within the household, in addition to gender, level of income and level of education (as in the CART analysis). In the specific case of the FKI, it is important to underline that targeting women as a single homogenous cluster in need of financial education is, again, a choice with potentially limited effects. Indeed, women with a higher level of educational attainment (university degree), who are married or cohabitating, and who are in the labor force earning medium- to high-level salaries, show, on average, degrees of financial knowledge that are similar to those attained by men. Those who are in greater need of receiving educational support on basic financial issues are women with limited education or those not in the labor force.

Considering the global financial literacy index proposed by the OECD INFE guidelines (see Fig. 5), the OLS regression identified gender, age, involvement in the financial decisions of the household, marital status, level of education, income and area of residence as relevant determinants of the individual's level of financial literacy. CART analysis restricted the number of relevant covariates, highlighting that the level of education was the first discriminatory variable to define individuals with lower and higher levels of financial literacy. Next, among individuals who had attained tertiary education, the level of income was the second most important discriminatory variable, which helped to identify a specific cluster with a high educational level but a low-income level. Once the level of income was considered, the area of residence was important for medium-income individuals, whereas gender became relevant for high-income individuals. On the other side of the "tree"—individuals with educational attainment up to the secondary level—the geographical area of residence, first, and age, second, were found to be discriminating factors. In summary, a financial education program targeting women as a homogenous (and vulnerable) cluster in need of educational support would not take into account the fact that only highly educated, high-income women are in need of such a specific program, whereas all other women can be addressed by finance programs targeting men with, for example, a low income and lower educational attainments.

So far, by critically reviewing the outcomes of one of the most comprehensive surveys conducted in Italy, we have shown that, according to the diverse statistical methods used to analyze the same financial literacy indexes, different insights about the level of financial literacy of individuals can be revealed, which have important implications for policies (including the design of education programs) aimed at improving this literacy.

A further step toward understanding a latent variable such as financial literacy could come from the adoption of statistical approaches that are able to provide information on the reliability and validity of the measures used. In the next section, we apply a well-known psychometric technique—item response theory (IRT)—in an area where such

techniques are not often applied; i.e., financial literacy measurement. To the best of our knowledge, only a few studies have explored the viability of these models for assessing financial literacy (Bongini et al. 2012, 2015; Knoll and Houtts 2012; Despard and Chowa 2014).

Methods

Sample

The proposed procedure for data handling was applied to a sample comprising 1247 Italian residents of at least 18 years of age who were reached via CATI. The sample was obtained by appropriate stratification across several dimensions (gender, age, geographical area, and municipality size).

Table 1 shows the distribution of the sample relative to the main sociodemographic variables. The respondents' average age is approximately 50 years. Almost 60% are married or cohabitants; and 42.3% are employed. The median family income declared by the respondents is approximately 1900 euros. Regarding the educational level of the respondents, approximately 21% received only primary education, 29% secondary education (lower level), 31% secondary education (upper level), and 11% tertiary education. In terms of the geographic composition of the sample, 46.5% of the respondents live in the northern region of the country; approximately 31.01% reside in small municipalities (up to 10,000 inhabitants); and 23.5% in large cities (above 100,000 inhabitants).

Statistical analysis: item response theory

The issue of the most appropriate way to measure literacy has attracted increasing attention in educational research over the last two decades. One important aim in measurement is to build tests with high validity and reliability. The two most popular frameworks in educational measurement are classical test theory (CTT) and item response theory (IRT) (Hambleton and Jones 1993). In general, CTT has dominated the area of standardized testing because of its weak assumptions and its easy interpretation. Indeed, the indexes proposed by the OECD approach and discussed above rely on CTT. Despite these features, CTT has been criticized since the score on a test is not an absolute characteristic of the respondent. In fact, it depends on the content of the test. Moreover, the difficulty of the items may vary depending on the sample of respondents who take a specific test. It is therefore difficult to compare the data of respondents between different tests. For these reasons, IRT was originally developed to overcome the problems with CTT.

The specific feature that makes IRT models increasingly popular in many areas of research is the presence of a metric that considers both the test's difficulty and the respondent's specific abilities. IRT aims to measure one or more ordinal/quantitative latent variables on a metric level of measurement, and it is fit to quantify aspects such as ability and personal traits. For these reasons, it has been widely adopted in educational research and psychometrics, where researchers develop and design exams, maintain banks of items for exams, and measure the items' difficulties for successive versions of exams by the use of IRT (Bond and Fox 2007; Goldstein 1979). For example, in computerized adaptive testing (CAT), the respondents respond to items that are optimally selected to assess their attitude or abilities. The respondents may receive no common

Table 1 Sample distribution

Variable	Categories	Frequency	%
Gender	Male	599	48.0
	Female	648	52.0
Age	18–24	112	9.0
	25–39	299	23.9
	40–64	531	42.6
	Over 65	305	24.5
Geographical area of residence	Northwest	237	19.0
	Northeast	343	27.5
	Central Italy	249	20.0
	Southern Italy and islands	418	33.5
Size of municipality of residence	Up to 10,000 inhabitants	387	31.0
	From 10,000 to 30,000 inhabitants	298	24.0
	From 30,000 to 100,000 inhabitants	268	21.5
	Above 100,000 inhabitants	293	23.5
Educational attainment	Primary school or lower	73	5.9
	Lower secondary school	292	23.4
	Upper secondary school	569	45.6
	High school or above	313	25.1
Job status	Employed	584	46.8
	Unemployed seeking employment	103	8.2
	Inactive not seeking employment	560	44.9
Marital status	Married/cohabitant	695	55.7
	Unmarried	370	29.7
	Divorced	82	6.6
	Widowed	96	7.7
Living with children	No	857	68.7
	Yes	390	31.3
Income (€)	≤ 1500	202	16.2
	[1501; 1900]	426	34.2
	[1901; 3300]	220	17.6
	≥ 3301	99	7.9
	Unknown	300	24.1

items. IRT helps to select the items for a respondent and to measure the scores across different subsets of items. For instance, several aptitude tests need IRT to estimate the abilities of the respondents, such as the Armed Services Vocational Aptitude Battery, the Scholastic Aptitude Test (SAT), and the Graduate Record Examination (GRE). Several individual intelligence tests adopt IRT to manage the tests, such as the Woodcock–Johnson Psycho-Educational Battery, the Differential Ability Scales, and the Stanford-Binet test (Embretson and Reise 2013). Furthermore, several researchers have applied IRT to personality trait measurements (Reise and Waller 1990), as well as to attitude measurements and behavioral ratings (Engelhard and Wilson 1996).

The Program for International Student Assessment (PISA) surveys has been adopting IRT models since 2000 (Liu et al. 2008). Moreover, personal properties or item characteristics can be included in IRT models to explain person or item effects, obtaining explanatory item-response models (De Boeck and Wilson 2004). Until very recently, the

analysis of financial literacy has relied only on CTT. To the best of our knowledge, only a few studies have used IRT in this domain (Knoll and Houts 2012; Bongini et al. 2012, 2015; Despard and Chowa 2014).

In general, IRT models convert raw scores into linear and reproducible measurements. An IRT model has two properties, which require checking in order to ensure the model's validity. Those properties are unidimensionality and local independence. The unidimensional property requires that the items of a questionnaire share a common primary construct (i.e., that they all measure financial literacy), while the local independence property requires that the items are significantly independent of each subpopulation of respondents whose members are homogeneous with respect to the latent trait measured (for instance, gender or race).

According to IRT models, an individual's response to an item is determined by his/her level of knowledge (alternatively, ability or trait) of the latent variable under investigation (e.g., financial literacy), and by the level of difficulty of the given item. IRT models define the score (number of items answered correctly) of a particular respondent as a probability function of his/her ability and item difficulty. One way of expressing IRT models is in terms of the probability that an individual with a particular trait will correctly answer an item that has a particular level of difficulty, as expressed in the following formula:

$$P(X_{pik} = 1 | \theta_p, \beta_{ik}) = \frac{e^{(\theta_p - \beta_{ik})}}{1 + e^{(\theta_p - \beta_{ik})}} \quad (1)$$

In Formula (1), X_{pik} refers to the response X made by the p -th individual to the i -th item (k refers to the possible level of the i -th item⁴); θ_p refers to the level of knowledge (ability) of financial literacy of the p -th individual; and β_{ik} is the level of financial literacy (difficulty) required to reach level k of the i -th item. In addition, we let β_i denote the average level of financial literacy for the i -th item.

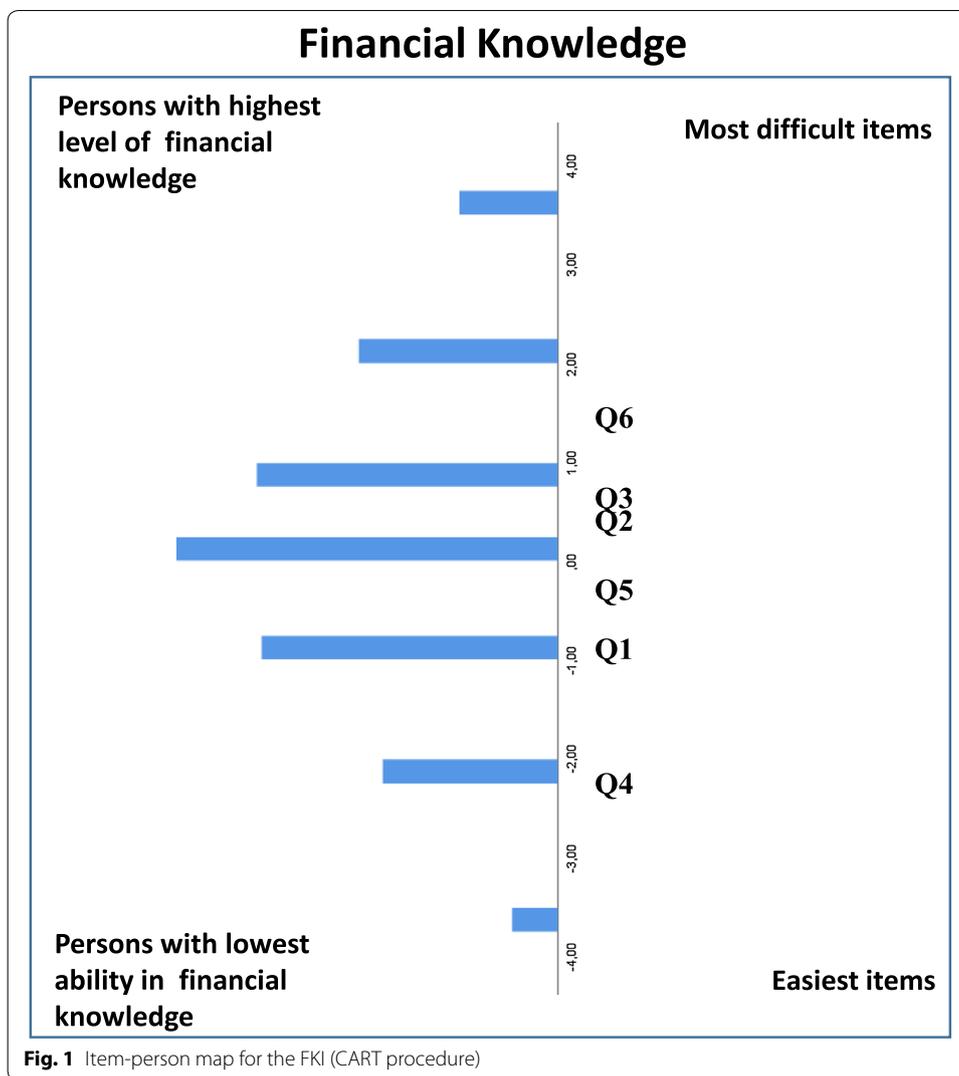
A typical representation of IRT is an "item map" where the item difficulties can be placed like points along a line and the person's ability as a point along the same line. In Fig. 1, we apply this method to the data underlying the FKI described in the previous section.

To answer our central research question—i.e., whether survey outcomes are sensitive to the data handling method employed—we applied IRT analysis to the survey data, checking the two properties of unidimensionality and local independence. Our aims are, firstly, to test whether the selected items were indeed measuring the same latent construct (i.e., an individual's financial knowledge, financial attitude, financial behavior, and level of financial literacy); and, secondly, to ensure the local independence property by assessing whether the instrument is measuring the specific object. Our third aim is to analyze the attributes of the items and the respondents on the same scale, via the item-person map, to convey easy-to-read information about the distribution of the respondents and the chosen items.

Results

Table 2 displays the misfit indexes (Wright 1999; Bond and Fox 2007) for our three elementary indexes (financial knowledge, financial attitude, and financial behavior) and for the overall latent variable of financial literacy. As the term implies, a misfit

⁴ The items have ordered categories $1, 2, \dots, k, \dots, K$, and K could vary among items.



is an observation that cannot fit into the overall structure of the questionnaire and is an indicator of how well the data conform to the IRT model parameters. In this work, we used the index based on the average value of the squared residuals (MNSQ). Two types of fit statistics are addressed by the MNSQ: infit (the weighted average of the squared residuals) and outfit (unweighted average of the squared residuals) (Bond and Fox 2007). Guidelines vary according to test, item and respondent characteristics, but for general purposes, an MNSQ value in the interval [0.5–1.5] means that the item is “productive” for the measurement. In contrast, for values greater than 2.0, the item is considered degrading for the measurement (Linacre 2006). Almost all the items used for computing the three sub-indexes are “productive for measurement”, which means that they do not distort the measure under investigation; i.e., they all measure the same latent construct. In the case of the FBI and the overall financial literacy index, the test confirmed that all but three items are coherent and finalized to measure the specific latent variable. However, such items did not degrade the measurement system; thus, they can be maintained in the questionnaire. In summary, we can confirm

Table 2 The measure of the items for FBI, FAI, FKI, and financial literacy index

z	Financial knowledge index		Outfit MNSQ		Financial attitude index		Outfit MNSQ		Financial behavior index		Outfit MNSQ		Financial global index		Outfit MNSQ	
	Measure	SE	INFIT	MNSQ	Measure	SE	INFIT	MNSQ	Measure	SE	INFIT	MNSQ	Measure	SE	INFIT	MNSQ
B1									-2.19	0.09	0.36	0.27	-0.99	0.09	0.26	0.23
B2									-1.93	0.09	0.49	0.41	-0.78	0.09	0.43	0.4
B3									-1.22	0.08	0.66	0.52	-0.18	0.08	0.74	0.68
B4									0.41	0.07	0.87	0.84	1.19	0.06	0.97	0.97
B5									1.87	0.07	0.99	1.07	2.41	0.06	0.99	0.99
B6									-1.18	0.08	0.86	0.87	-0.13	0.08	0.78	0.73
B7									1.55	0.07	1.62	1.76	2.13	0.06	1.82	1.77
B8									2.69	0.08	1.12	1.15	3.09	0.07	1.07	1.13
A1					-0.65	0.03	1.21	1.06					-5.87	0.03	1.05	0.9
A2					-0.45	0.03	0.84	0.78					-5.73	0.03	0.9	0.88
A3					1.1	0.03	0.94	1.16					-4.67	0.03	0.95	1.09
K1	-0.99	0.07	1.00	1.02									0.88	0.06	0.96	0.94
K2	0.63	0.07	0.87	0.82									2.05	0.06	0.94	0.93
K3	0.81	0.07	0.99	1.00									2.17	0.06	0.97	0.97
K4	-2.01	0.09	1.07	1.3									0.18	0.07	0.9	0.88
K5	-0.16	0.07	0.93	0.91									1.45	0.06	0.98	0.98
K6	1.72	0.08	1.14	1.31									2.81	0.07	1.00	1.02

Table 3 Standardized residual variance of the indexes

	Financial behavior index		Financial attitude index		Financial knowledge index		Financial global index	
	Empirical	Modeled	Empirical	Modeled	Empirical	Modeled	Empirical	Modeled
Raw variance explained by the measures (%)	53.4	55.8	59.4	58.9	48.2	50.8	80.7	80.1
Unexplained variance in 1st contrast (%)	9.8		7.4		8.4		2.1	

that the items proposed in the OECD/INFE questionnaire are indeed good measures of one latent variable and can be used together.

A second relevant property that needs to be assessed is local independence, which ensures that the instrument is measuring the specific object. For this purpose, we apply principal component analysis to the standardized residuals (Smith 2002). Table 3 shows the standardized residual variance decomposition for our set of indexes.

The raw variances of the empirical model explained by each index closely match the expected raw variances (modeled). Moreover, because the modeled values for the three indexes are in the interval [50–60%], the measurement scales can be considered fairly good. Regarding the overall financial literacy index, which exhibits a value greater than 80%, the measurement scale is considered excellent. Furthermore, for the three indexes, the unexplained variances in the 1st contrast demonstrate that the instrument is good (Fisher 2007) since it falls within the required interval [5–10%]. Given that the value of unexplained variances for the global financial literacy index is less than 3%, the measurement instrument is excellent. In summary, we confirm, on solid statistical grounds, that the items used to build the four indexes meet the required unidimensional and local independence traits and are appropriate to define the level of financial literacy of an individual.

The third aim of our analysis was to assess the attributes of the items and respondents at the same time via the item-person map. Figure 1 presents the item-person map for the FKI. Maps produced by IRT models can be used to quickly communicate complex information and do so in a presentational format that can be easily understood. Indeed, if we were not using an IRT metric, we would have been unable to measure, on the very same scale, both the respondents' ability and the questions' difficulty. In fact, in the case of the financial knowledge items, the item difficulty scores ranged between 0 and 1247 (i.e., the whole sample): zero was applied to the case where no respondent answered each question correctly, and 1247 applies to the case where the whole sample answered each question correctly. Conversely, a person's ability is measured on a scale ranging from 0 to 6: 0 corresponds to a person who was unable to answer any item correctly, and 6 applies to a person who answered the whole set of questions correctly. Therefore, the two metrics are not directly comparable.

The IRT item-person map shown in Fig. 1 orders the level of financial knowledge of the respondent (left-hand side), and the difficulty of the multiple-choice questions (right-hand side). The questions at the top of the scale were more difficult to answer; hence the test becomes easier further down the scale. The individuals with the least financial ability (at

Table 4 Percentage of correct answers for the six items of the FKI

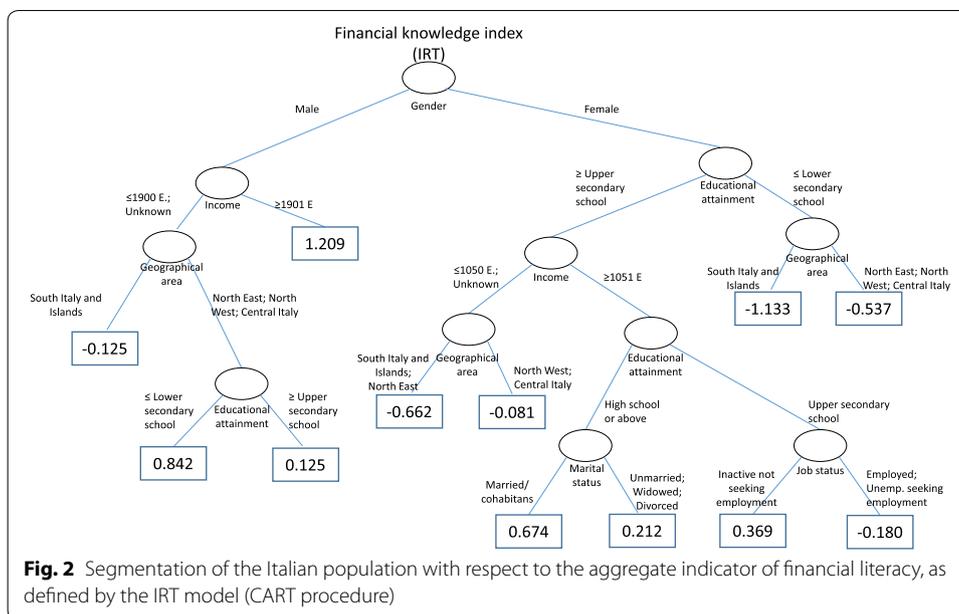
Item	% of correct answers
K1	68.6
K2	41.6
K3	39.4
K4	82.6
K5	55.0
K6	25.7

the bottom of the scale) had difficulty even with the easiest concepts (e.g., the relationship between risk and return); whereas the individuals with the most financial literacy (at the top of the scale) had no difficulty performing any of the activities implied by the questions. In particular, the respondents on the upper left-hand side were said to be “better” or “smarter” than the items on the lower right-hand side, which means that these easier items were not difficult enough to challenge highly proficient individuals. On the other hand, the items on the upper right-hand side outsmarted the individuals on the lower left-hand side, which implies that these difficult items were beyond the level of ability possessed by our sample. Items 4 and 6 were the easiest and most difficult to answer, respectively. This conclusion is also supported by the frequency distribution of the answers given to the six items concerning the construct of financial knowledge. Table 4 lists the percentage of correct answers for the six items.

The relevant contribution of IRT lies in the fact that the map reproduces directly the frequency distribution of the respondents with respect to their financial knowledge (ability) and the position of the items with respect to their difficulty in the financial knowledge construct. The unit of measurement of difficulty and ability is the same. For instance, item 4, with a difficulty equal to -0.99 , was correctly answered by 85.3% of the respondents. This is equivalent to saying, ‘85.3% is the proportion of respondents who have an ability greater than -0.99 ’.

Having confirmed that the items were correctly chosen, and having investigated the relationship between difficulty and ability, a researcher is subsequently provided with a number of statistical methods to further investigate the socioeconomic characteristics of the respondents in relation to the IRT measure. For instance, it might be useful to evaluate whether a specific subgroup (defined by age, gender, or education) is disadvantaged or advantaged with respect to single items (numeracy problems, behavioral aspects, or attitude issues) and the whole issue under investigation (financial literacy). Differential item functioning (DIF) is a method that can uncover such differences, as explored by Bongini et al. (2012, 2015), who found a gender gap among university students on a single item (but not one that referred to the whole construct of financial literacy). Alternatively, one can include IRT measures and the socioeconomic characteristics of the respondents in a latent regression model, which provides a powerful framework to detect and analyze group differences that considers the characteristics of both items and individuals simultaneously (De Boeck and Wilson 2004).

Finally, CART analysis can be applied to the IRT measure. In this study, we applied CART analysis to the overall financial literacy index to compare the results when applied to the same construct (financial literacy) but measured through two different methods,



CTT (Fig. 5) and IRT (Fig. 2). It is immediately apparent that the same approach applied to two different ways of constructing the same latent variable delivers different results with respect to relevant clusters differing in their level of financial literacy. In other words, depending on how we handle financial literacy data, through CTT or IRT models, we end up with dissimilar outcomes about who needs more financial education.

Discussion/conclusion

The present paper aimed to provide insight in order to improve the procedures for analyzing data that describe a latent variable such as financial literacy, leveraging the recent Italian national survey based on the approach proposed by the OECD through its INFE. As underlined in the introduction, assessing the baseline level of financial literacy represents an indispensable prerequisite to the design of effective education programs; that is, interventions that successfully address specific target groups with particular educational needs. The evidence provided in this paper shows, firstly, that different methods of analysis applied to the same measure of financial literacy deliver different results; and, secondly, that the same method of analysis applied to different measures of financial literacy also delivers different results. Consequently, we can state that the method of data analysis is crucial for the subsequent step of devising successful education programs in the field of financial literacy among different target groups. In particular, our findings show that adopting a specific method of data analysis delivers results that would not be obtained by adopting an alternative method, thus indicating that different approaches cannot be considered interchangeable. These findings suggest further improvements to the process of financial literacy evaluation which we summarize here.

First, CTT has long been proven to be outdated as regards defining people’s level of financial literacy. A basic test should be integrated into more sophisticated models where the difficulty of the items and the ability of the respondents are considered. From this perspective, using IRT helps to define for every possible test item difficulty the existence of a weighted score that corresponds to that level of ability, opinion, or feeling of

the respondents. Moreover, when the assumptions of IRT are proven, its estimates of the item parameters are independent of the sample. A respondent should show the same ability, independent of the set of items adopted; and conversely, a given item should have the same difficulty, independent of the respondents.

Second, applying alternative and more sophisticated methods of data analysis to financial literacy data enables researchers to target specific population groups. Instead of assuming that sharing the same personal characteristics among individuals (e.g., gender) necessarily means sharing the same financial literacy needs, the results of our CART analysis suggest that women should not be considered a homogeneous group in terms of their level of financial literacy. Consequently, policy makers cannot treat “women” as a potentially identical target of the same education program; rather, they should differentiate and develop specific programs depending on the different cluster to which women belong (e.g., educational level, residential area). In this regard, with the goal of targeting people (especially the more financially vulnerable ones) in an ever more detailed and precise way, the data analysis methods used in this study might offer the possibility to also include other individual non-cognitive characteristics such as personality traits, which were recently proven to be a fundamental aspect of financial behavior. For example, research has investigated conscientiousness (Roa et al. 2017) and impulsivity (Baldi et al. 2013; Iannello et al. 2015; Bongini et al. 2015), as well as the social roles of respondents, such as homemakers vs. financial workers (Croson and Gneezy 2009; Dwivedi et al. 2015).

Many of the studies carried out in Italy to date have focused on financial knowledge (the cognitive aspect of financial literacy). However, future research should perhaps carry out more in-depth analysis of soft skills rather than content knowledge, such as the confidence to be proactive, and a willingness to take investment risks. For example, in a meta-analysis carried out by Fernandes et al. (2014), measured knowledge of financial facts had a weak relationship to financial behavior in econometric studies, controlling for omitted variable bias. As pointed out by many authors (e.g., Worthington 2006; Nicolini 2017), financial literacy should be tested against an individual’s needs and the context in which they live, not against a large set of available financial products and services, since consumers will never need or use most of these products and services. The assumption here is that an individual’s financial literacy should not be measured in “a linear sense” but, rather, with respect to the set of knowledge that is necessary to deal with specific financial needs, desires, expectations and fears of specific groups of consumers. From this perspective, “financial literacy becomes a multidimensional construct, with an individual being knowledgeable in certain domains (e.g., investing) while showing a deep lack of knowledge in others (e.g., borrowing)” (Nicolini 2017, p. 35). However, a lack of knowledge in a specific area is not considered a very critical gap if the individual is not called to make financial decisions in that area.

In line with this view, we are aware that the questionnaire used in our study involves some aspects that warrant critique. That questionnaire was a version of the OECD/INFE questionnaire tailored to a national survey and, as pointed out by Robson and Splinter (2015), one problem with national surveys is that there is no clear way to assess individual responses and micro-level changes over time in regard to behavior. As for studies that use the questionnaire to provide reliable information on what people do in the financial domain, another limitation is that the questionnaire tests behavior with self-assessed questions that deal with financial problems and tasks which may be not

be realistic for every respondent. In fact, it focuses primarily on one aspect of an individual's financial capability,⁵ attaching less importance to the context. Further research should take into account social and contextual issues, as suggested by some institutions that promote financial inclusion and financial wellbeing (e.g., CYFI 2012; CFPB 2015), and by authors who are critical of mainstream approaches to financial literacy and work with people living on low incomes (e.g., Landvøgt 2006; Rinaldi 2016).

To conclude, our research suggests that financial literacy research should be open to new and alternative approaches to measurement, while being aware that different data analysis methods can produce different results. Therefore, different types of analysis are called for. Additionally, researchers should be clear about why one method is to be preferred to another, and why one set of results are more useful than another set. This sort of information would be useful to policy makers who are keen to design more efficient and more effective financial education programs for target groups.

Abbreviations

CAT: computerized adaptive testing (cat); CART: classification and regression tree; CCT: classical test theory; CFPB: Consumer Financial Protection Bureau; CYFI: Child Youth Financial International; DIF: differential item functioning; FAI: financial attitude index; FBI: financial behavior index; FFI: financial familiarity index; FKI: financial knowledge index; FPI: financial planning index; GRE: Graduate Record Examination; INFE: International Network on Financial Education; IRT: item response theory; MNSQ: average value of the squared residuals (mean square); OLS: ordinary least square; PISA: Program for International Student Assessment; SAT: Scholastic Aptitude Test.

Authors' contributions

Conceived and designed the research: PB, MZ, ER. Formal analysis: PB, MZ. Analyzed the data: PB, MZ. Methodology: PB, MZ. Data interpretation: PB, PI, MZ, ER. Writing—original draft preparation: PB, PI, MZ, ER. Writing—review & editing PB, AA. Supervision: PB, AA. All authors read and approved the final manuscript.

Author details

¹ Department of Business and Law, Università degli Studi di Milano-Bicocca, Milan, Italy. ² Department of Psychology, Università Cattolica del Sacro Cuore, Largo Gemelli, 1, 20123 Milan, Italy. ³ Department of Statistics and Quantitative Methods, Università degli Studi di Milano-Bicocca, Milan, Italy.

Acknowledgements

Not applicable.

Competing interests

The authors declare that they have no competing interests.

Availability of data and materials

Data will not be shared, but the dataset supporting the conclusions of this article can be requested at Pattichiarì (now Fondazione per l'Educazione Finanziaria e al Risparmio - info@feduf.it).

Consent for publication

Not applicable.

Ethics approval and consent to participate

Authors did not collect data themselves, but they used the dataset "Rilevazione ICF: Indice di Cultura Economico-Finanziaria" (™ Consorzio PattiChiarì 2012).

Funding

Consorzio Pattichiarì (now Fondazione per l'Educazione Finanziaria e al Risparmio—FEDUF) funded the collection of the data. No funds were received for design, analysis and interpretation of the data of the present paper.

Appendix 1

See Tables 5, 6 and Figs. 3, 4, 5.

⁵ A broader concept that can be defined as the internal capacity to act in one's best financial interest, given socioeconomic environmental conditions—The World Bank 2013.

Table 5 Results of the ordered probit regressions applied to each elementary i. Source: Baglioni et al. (2018)

	Financial familiarity index (FFI)			Financial behavior index (FBI)			Financial attitude index (FAI)			Financial knowledge index (FKI)			Financial planning index (FPI)		
	I	II	III	I	II	III	I	II	III	I	II	III	I	II	III
Individual and household characteristics															
Female	-0.472*** (0.088)	-0.404*** (0.063)	-0.358*** (0.108)	-0.007 (0.077)	-0.031 (0.059)	-0.071 (0.107)	0.118 (0.085)	0.087 (0.063)	0.091 (0.106)	-0.542*** (0.083)	-0.641*** (0.062)	-0.650*** (0.103)	-0.173** (0.083)	-0.149** (0.062)	0.029 (0.110)
Age	0.093*** (0.017)	0.098*** (0.012)	0.083*** (0.014)	-0.011 (0.014)	0.006 (0.012)	-0.007 (0.012)	0.030* (0.017)	0.031*** (0.013)	0.012 (0.014)	0.041*** (0.016)	0.032*** (0.011)	0.022** (0.013)	-0.019 (0.017)	0.004 (0.013)	-0.002 (0.014)
Age ²	-0.0009*** (0.0001)	-0.0008*** (0.0001)	-0.0008*** (0.0001)	-0.0000 (0.0001)	-0.0001 (0.0001)	-0.0000 (0.0001)	-0.0003* (0.0002)	-0.0003*** (0.0001)	-0.0002 (0.0001)	-0.0004** (0.0001)	-0.0002** (0.0001)	-0.0002** (0.0001)	0.0002 (0.0002)	-0.000 (0.0001)	-0.0000 (0.0001)
Nr. household members	0.070 (0.044)	-0.010 (0.030)	-0.044 (0.032)	-0.018 (0.044)	-0.038 (0.027)	-0.021 (0.029)	0.017 (0.044)	0.013 (0.030)	0.027 (0.032)	0.032 (0.045)	0.011 (0.030)	0.051 (0.032)	-0.075 (0.048)	-0.076** (0.032)	-0.058* (0.033)
Nr. underage children	0.003 (0.060)	0.099** (0.047)	-0.013 (0.052)	0.066 (0.062)	0.099** (0.042)	0.054 (0.046)	0.058 (0.061)	0.051 (0.044)	-0.031 (0.048)	0.018 (0.061)	0.065 (0.045)	-0.003 (0.048)	0.147** (0.061)	0.104** (0.045)	0.067 (0.048)
Participation in financial decisions	0.245*** (0.071)		0.251*** (0.054)	0.092 (0.066)		0.170*** (0.048)	0.139** (0.073)		0.151*** (0.055)	0.134** (0.067)		0.163*** (0.051)	0.072 (0.068)		0.111** (0.052)
Married/cohabitant			0.444*** (0.118)			0.184* (0.106)			0.397*** (0.108)			0.205* (0.107)			0.313*** (0.113)
Female*married/cohabitant			-0.120 (0.134)			0.035 (0.129)			-0.038 (0.132)			-0.008 (0.126)			-0.297** (0.135)
Education															
Educational level	0.375*** (0.086)	0.400*** (0.047)	0.379*** (0.061)	-0.037 (0.083)	0.149*** (0.045)	0.039 (0.059)	0.131 (0.090)	0.168*** (0.049)	0.124** (0.062)	0.406*** (0.090)	0.399*** (0.048)	0.385*** (0.064)	0.050 (0.087)	0.207*** (0.048)	0.119** (0.061)
High school type	0.002 (0.070)		0.098* (0.051)	0.125* (0.070)		0.178*** (0.055)	0.154** (0.073)		0.085* (0.051)	0.014 (0.074)		0.068 (0.054)	0.141** (0.072)		0.133** (0.054)
Vocational school (business)		0.410*** (0.091)			0.107 (0.081)			0.038 (0.092)			0.387*** (0.090)		0.087 (0.086)		
Parents' educational level	-0.040 (0.041)	0.060** (0.029)	0.038 (0.031)	0.005 (0.041)	-0.017 (0.029)	-0.030 (0.030)	-0.147*** (0.043)	-0.066** (0.033)	-0.073* (0.033)	-0.011 (0.041)	0.096*** (0.033)	0.075** (0.033)	-0.039 (0.043)	0.025 (0.032)	0.014 (0.033)

Table 5 (continued)

	Financial familiarity index (FFI)			Financial behavior index (FBI)			Financial attitude index (FAI)			Financial knowledge index (FKI)			Financial planning index (FPI)		
	I	II	III	I	II	III	I	II	III	I	II	III	I	II	III
Labor force status and income															
Labor force participation	0.128 (0.107)	0.271*** (0.079)	0.195*** (0.081)	0.042 (0.101)	0.081 (0.075)	0.048 (0.077)	-0.176 (0.110)	-0.145* (0.079)	-0.184** (0.082)	-0.112 (0.098)	-0.095 (0.075)	-0.152** (0.076)	0.196* (0.107)	0.247*** (0.077)	0.190** (0.079)
Family income	0.133*** (0.019)			0.049*** (0.017)			0.028 (0.019)						0.089*** (0.019)		
Geographical variables															
Town size	-0.085 (0.060)			-0.006 (0.055)			-0.081 (0.061)						-0.108 (0.072)		
Northwest	0.286** (0.117)			0.108 (0.112)			0.130 (0.121)						0.174 (0.120)		0.229** (0.093)
Northeast	0.345*** (0.124)			0.186* (0.114)			0.162 (0.134)						0.289** (0.119)		0.336*** (0.096)
South	-0.307*** (0.108)			0.070 (0.109)			-0.108 (0.119)						0.025 (0.113)		-0.039 (0.092)
Nr. Obs.	719	1247	1247	719	1247	1247	719	1247	1247	719	1247	1247	719	1247	1247
Pseudo R-sq.	0.151	0.086	0.114	0.014	0.009	0.016	0.030	0.009	0.022	0.076	0.058	0.066	0.050	0.021	0.033
Wald Chi-sq. [p value]	255.28	260.61	296.39	39.41	44.39	73.99	51.07	29.16	64.81	201.04	243.12	288.92	91.28	68.96	104.98
Log Pseudo-L	-793.078	-1486.64	-1440.98	-1.231.241	-2.195.900	-2.180.538	-844.505	-1.476.702	-1.457.225	-1191.22	-2109.85	-2092.90	-869.94	-1548.69	-1.528.917

Ordered probit MLE; asymptotically robust SE in parenthesis

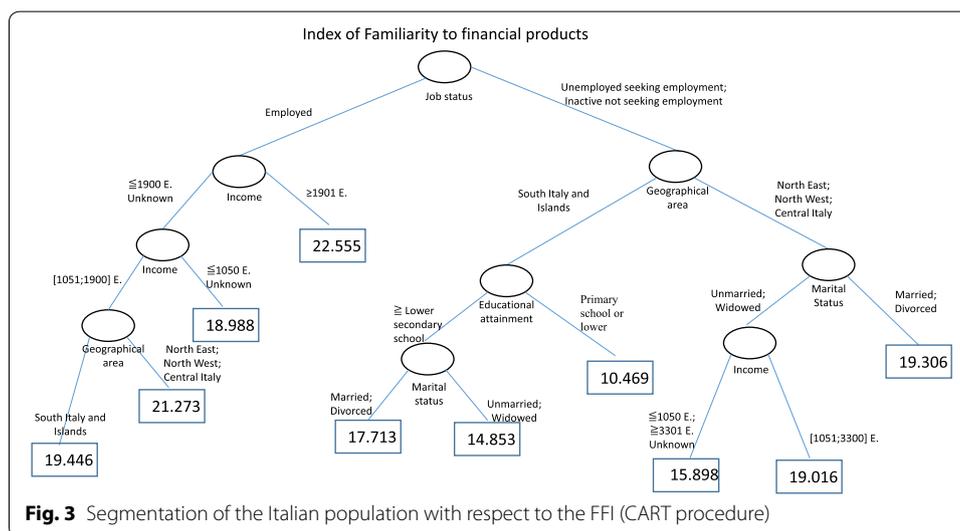
Significance levels: *** 1%; ** 5%; * 10%

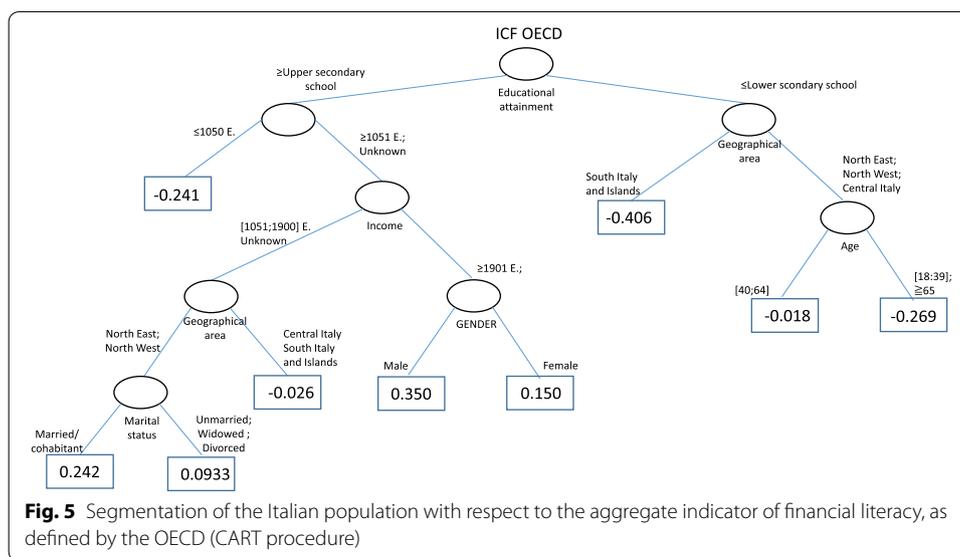
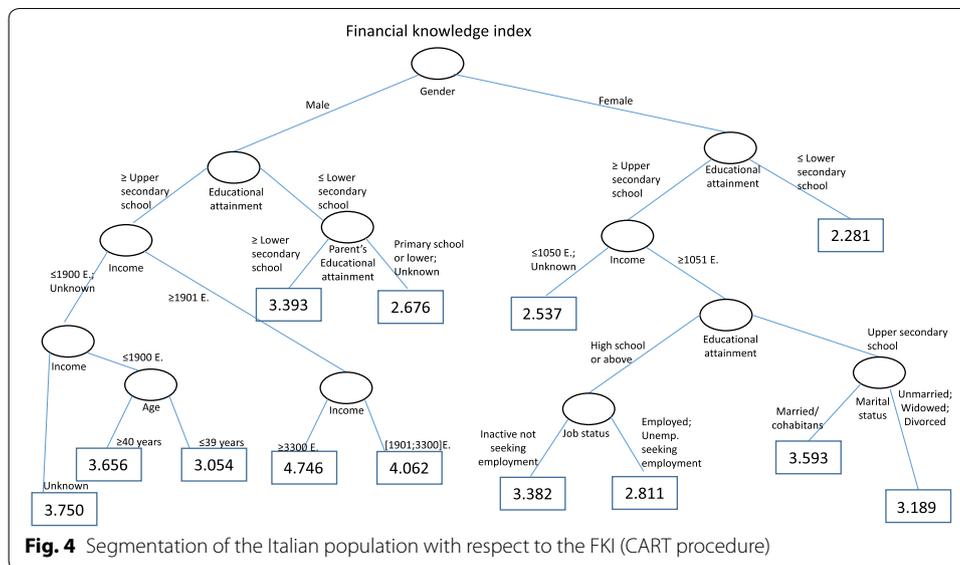
Table 6 Results of an OLS regression applied to the aggregate indicators of financial literacy. Source: Baglioni et al. (2018)

	Global financial literacy index		
	I	II	III
Female	-0.163*** (0.040)	-0.197*** (0.032)	-0.158*** (0.056)
Age	0.020*** (0.008)	0.030*** (0.006)	0.018*** (0.002)
Age ²	-0.0002*** (0.000)	-0.0003*** (0.000)	-0.0002*** (0.000)
Nr. household members	0.006 (0.023)	-0.019 (0.016)	-0.007 (0.016)
Nr. underage children	0.047 (0.030)	0.075*** (0.023)	0.015 (0.024)
Participation in fin. decisions	0.113*** (0.032)		0.147*** (0.025)
Married/cohabitant			0.265*** (0.059)
Female*married/cohabitant			-0.077 (0.069)
Educational level	0.149*** (0.042)	0.235*** (0.025)	0.183*** (0.031)
High school type	0.076** (0.035)		0.097*** (0.027)
Vocational school (business)		0.187*** (0.045)	
Parents' educational level	-0.038* (0.021)	-0.013 (0.017)	0.000 (0.016)
Labor force participation	0.033 (0.052)	0.069* (0.041)	0.021 (0.041)
Family income	0.067*** (0.009)		
Town size	-0.033 (0.028)		
Northwest	0.160*** (0.058)		0.110** (0.046)
Northeast	0.205*** (0.059)		0.201*** (0.047)
South	-0.070 (0.056)		-0.134*** (0.045)
Constant	-1.000*** (0.218)	-0.935*** (0.159)	-0.896*** (0.165)
Nr. Obs.	703	1221	1221
R-sq.	0.302	0.168	0.238
Model F	19.36	29.07	24.36
[p-value]	[0.000]	[0.000]	[0.000]

OLS regression; asymptotically robust SE in parenthesis

Significance levels: *** 1%; ** 5%; * 10%





Publisher’s Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Received: 7 February 2018 Accepted: 13 November 2018

Published online: 21 November 2018

References

Abi-PattiChiari (2014) *Le competenze economico-finanziarie degli italiani*. Bancaria Editrice, Roma

Atkinson A, Messy F (2012) *Measuring financial literacy: results of the OCSE/International Network on Financial Education (INFE) Pilot Study*, OCSE Working Papers on Finance, Insurance and Private Pensions

Baglioni A, Colombo L, Piccirilli G (2018) On the anatomy of financial literacy in Italy. *Econ Notes* 47(2–3):245–304

Baldi PL, Iannello P, Riva S, Antonietti A (2013) Cognitive reflection and socially biased decisions. *Stud Psychol* 55:265–271

Bond TG, Fox CM (2007) *Applying the rash model. fundamental measurement in human sciences*. Psychology Press, London

- Bongini P, Trivellato P, Zenga M (2012) Measuring financial literacy among students: an application of rasch analysis. *Elect J Appl Stat Anal* 5(3):425–430
- Bongini P, Trivellato P, Zenga M (2015) Business students and financial literacy: when will the gender gap fade away? *J Financ Manag Markets Instit* 3(1):13–19
- Breiman L, Friedman J, Stone CJ, Olshen RA (1984) Classification and regression trees. Chapman & Hall/CRC, Boca Raton
- CFPB (Consumer Financial Protection Bureau) (2015) Financial well-being: the goal of financial education, Report. https://files.consumerfinance.gov/f/201501_cfpb_report_financial-well-being.pdf
- Coppola M, Langley G, Sabatini M, Wolf R (2017) When will the penny drop money, financial literacy and risk in digital age, Allianz International Pension Papers 1, Munich, Germany. http://gflrec.org/wpcontent/uploads/2017/01/AGIIP_P_117_FinancialLiteracy_FINAL.pdf?x87657
- Crosno R, Gneezy U (2009) Gender differences in preferences. *J Econ Lit* 47(2):1–27
- Child and Youth Finance International (2012) Children and youth as economic citizens: review of research on financial capability, financial inclusion, and financial education. Research Working Group Report. CYFI, Amsterdam. http://childfinanceinternational.org/index.php?option=com_mtree%26task=att_download%26link_id=374%26cf_id=200
- De Boeck P, Wilson M (2004) Explanatory item response models: a generalized linear and non linear approach. Springer, New York
- Despard MR, Chowa GS (2014) Testing a measurement model of financial capability among youth in Ghana. *J Consum Aff* 48(2):301–322
- Dwivedi M, Purohit H, Mehta D (2015) Improving financial literacy among women: the role of universities. *Economic Challenger*, October–December 2015. Retrieved from <https://ssrn.com/abstract=2668034>
- Embretson SE, Reise SP (2013) Item response theory. Psychology Press, Hove
- Engelhard G, Wilson M (1996) Objective measurement: theory into practice, vol 3. Ablex, Norwood
- Fernandes D, Lynch JG Jr, Netemeyer RG (2014) Financial literacy, financial education and downstream financial behaviors. *Manag Sci* 60(8):1861–1883
- Fisher WP (2007) Rating scale instrument quality criteria. *Rasch Meas Trans* 21:1095
- Goldstein H (1979) Consequences of using the rasch model for educational assessment. *Br Educ Res J* 5(2):211–220
- Grifoni A, Messy F (2012) Current status of national strategies for financial education: a comparative analysis and relevant practices, OCSE Working Papers on insurance and private pensions
- Hambleton RC, Jones RW (1993) Comparison of classical test theory and item response theory and their applications to test development. *Educ Meas Issue Pract* 12(3):38–47
- Hilgert M, Hogarth J, Beverley S (2003) Household financial management: the connection between knowledge and behavior. Technical Report 309–322. *Fed Res Bull*
- Huston SJ (2010) Measuring financial literacy. *J Consum Aff* 44:296–316
- Iannello P, Bionassi F, Nelli B, Zugno E, Colombo B (2015) The influence of menstrual cycle and impulsivity on risk-taking behaviour. *Neuropsychol Trends* 17:47–52
- Kempson E (2009) Framework for the development of financial literacy baseline surveys. OECD Working Papers on finance, insurance and private pensions, No. 1. OECD Publishing, Paris. <https://doi.org/10.1787/5kmdmdpz7m9zq-en>
- Knoll MAZ, Houts CR (2012) The financial knowledge scale: an application of item response theory to the assessment of financial literacy. *J Consum Aff* 46(3):381–410
- Landvogt K (2006) Critical financial capability. In: Paper presented at the financial literacy, banking and identity conference, RMIT University, 25–26. <http://mams.rmit.edu.au/cwt46mna0lofz.pdf>. Accessed 20 Sept 2015
- Linacre JM (2006) Data variance explained by measures. *Rasch Meas Trans* 20:1045–1047
- Linciano N, Gentile M, Soccorso P (2017) Report on financial investments of Italian households. Commissione Nazionale per le Società e la Borsa (Consob). <http://www.consob.it/documents/46180/46181/rf2017.pdf/50db8f22-ee9f-4bd4-83ea-55224b820d9d>
- Liu OL, Wilson M, Paek I (2008) A multidimensional Rasch analysis of gender differences in PISA mathematics. *J App Meas* 9(1):18–35
- Lusardi A, Mitchell OS (2011) Financial literacy around the world: an overview. *J Pension Econ Financ* 10(4):497–508
- Lusardi A, Mitchell OS (2014) The economic importance of financial literacy: theory and evidence. *J Econ Lit* 52(1):5–44
- Mandell L (2007) Financial literacy of high school students. In: Xiao JJ (ed) *Handbook of consumer finance research*. Springer, New York, NY, pp 163–183
- Moore D (2003) Survey of financial literacy in Washington State: Knowledge, behavior, attitudes, and experiences. Technical Report n. 03–39, Social and Economic Sciences Research Center, Washington State University
- Nicolini G (2017) The assessment methodologies of financial literacy. In: Linciano N, Soccorso P (eds) *Challenges in ensuring financial competencies. Essays on how to measure financial knowledge, target beneficiaries and deliver educational programme*, Quaderni di finanza CONSOB, n.84, ottobre. Tiburtini s.r.l., Roma, pp 34–43
- OECD INFE (2011) Measuring financial literature: questionnaire and guidance notes for conducting an internationally comparable survey of financial literacy. OECD, Paris
- Reise SP, Waller NG (1990) Fitting the two-parameter model to personality data. *App Psychol Meas* 14:45–58. <https://doi.org/10.1177/014662169001400105>
- Remund DL (2010) Financial literacy explication: the case for a clearer definition in an increasingly complex economy. *J Consum Aff* 44(2):276–295
- Rinaldi E (2016) The relationship between financial education and society: a sociological perspective. *Italian J Social Educ* 8(3):126–148
- Roa MJ, Garron I, Barboza J (2017) The importance of numerical abilities, conscientiousness and financial literacy in financial decision-making: an empirical analysis in the Andean Region. In: Paper presented at the 3rd Cherry Blossom financial education institute April 6–7. Washington, DC
- Robson J, Splinter J (2015) A new (and better) way to measure individual financial capability, research report, prepared under contract to Vancity Credit Union. Carleton University, Ottawa. <http://carleton.ca/politicalmanagement/people/jenniferrobson>

- Smith EV Jr (2002) Detecting and evaluating the impact of multidimensionality using item fit statistics and principal component analysis of residuals. *J App Meas* 3(2):205–231
- The European House-Ambrosetti, Consorzio PattiChiari (2007) *L'Educazione Finanziaria in Italia. Riflessioni e proposte per migliorare l'educazione finanziaria del Paese*
- The World Bank (2013) Making sense of financial capability surveys around the world. A review of existing financial capability and literacy measurement instruments. Working paper, World Bank, Washington, DC. <http://responsiblefinance.worldbank.org/~media/GIAWB/FL/Documents/Misc/Financial-Capability-Review.pdf>
- Van Rooij M, Lusardi A, Alessie R (2011) Financial literacy and stock market participation. *J Financ Econ* 101(2):449–472
- Worthington A (2006) Predicting financial literacy in Australia. *Financ Serv Rev* 15:59–79
- Wright BD (1999) Fundamental measurement for psychology. In: Embretson SE, Hershberger SL (eds) *The new rules of measurement: what every psychologist and educator should know*. Erlbaum, Mahwah, pp 65–104

Submit your manuscript to a SpringerOpen[®] journal and benefit from:

- ▶ Convenient online submission
- ▶ Rigorous peer review
- ▶ Open access: articles freely available online
- ▶ High visibility within the field
- ▶ Retaining the copyright to your article

Submit your next manuscript at ▶ springeropen.com
