Financial literacy and stock market participation:
The moderating effect of country-specific social connectedness

University of Groningen
Faculty of Economics and Business
MSc International Financial Management

Uppsala University
Department of Business Studies
MSc Business and Economics

B

Luuk Arts
S2353539

Supervisor: dr. M.M. Kramer
Co-assessor: dr. R.O.S. Zaal
January 2018

Abstract

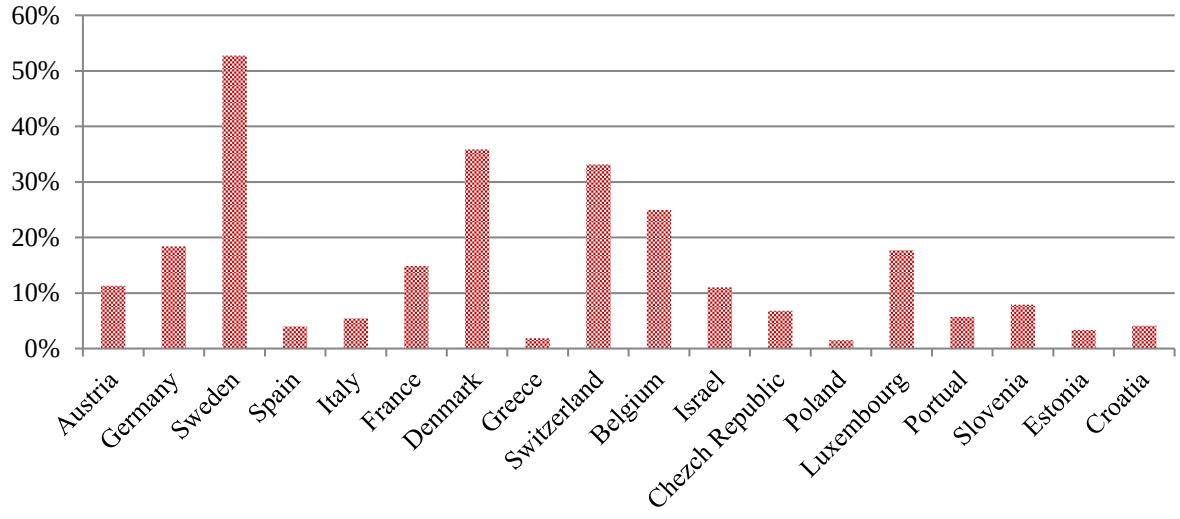
This research studies the moderating effect of country-specific social connectedness on the relation between financial literacy and stock market participation. This is done by using the extensive and multi-country SHARE data. The positive relation between financial literacy and stock market participation is reconfirmed. Moreover, the findings show that country-specific social connectedness significantly moderates the relation between financial literacy and stock market participation. The findings are robust and indicate that the predictive power of financial literacy on stock market participation decreases if country-specific social connectedness increases. This research is following up on contemporary literature and contributes to the explanation of the stock market participation puzzle on a macroeconomic scale.

Keywords: *stock market participation, financial literacy, country-specific social connectedness*
JEL-classification: *G11, G41, D83*

1. Introduction

Stock market participation is a phenomenon widely discussed in finance related activities and academic literature, while the relatively large differences in country-specific stock market participation rates have been left largely unexplored. Cocco, Gomes, and Maenhout (2005) show in their research that avoiding welfare loss can and should be achieved by investing in risky assets. This can be achieved for instance by participating in the stock market. A comprehensive set of traits explains the level of investments by means of stock market participation although stock market literacy takes on a prevailing role according to Balloch, Nicolae, and Philip (2014). Financial decision making is affected by an individual's level of financial literacy since individuals labeled as low literate are less likely to invest in stocks and therefore are less likely to participate in the stock market (Van Rooij, Lusardi, and Alessie, 2011). Yet in recent years, market liberalizations and structural reforms in pensions and social security plans have induced individuals to become more responsible for their own financial prosperity (Van Rooij et al., 2011). Although many scholars elaborate on the relevance of financial literacy associated to stock market participation and the importance of the latter, participation rates still are considerably low. Additionally, substantial differences exist in stock market participation rates between European countries (Guiso, Haliassos, and Jappelli, 2003). These two phenomena are generally considered as the stock market participation puzzle. Fig. 1 provides an overview of the differences in stock market participation rates between countries.

Figure 1
Stock market participation rate per European country. This figure provides an overview of the stock market participation as fraction of the population per country. Stock market participation consist of both direct and indirect stockholdings. Numbers are derived from Wave 6 of the Survey of Health, Ageing and Retirement in Europe (SHARE) database.



In addition to financial literacy, academic literature discusses other individual characteristics influencing stock market participation, which are age, gender, wealth, risk aversion, and education (see, e.g., Hong, Kubik, and Stein, 2004; Georgarakos and Pasini, 2011; Almenberg and Dreber, 2015). Georgarakos and Pasini (2011) state in their research that the average household's wealth holdings do not differ substantially between European countries, however, as shown in Fig. 1, the stock market participation rates do. They differ even to such extent that households with below average net wealth holdings in Sweden, Denmark, and Switzerland - countries with high stock market participation rates - show participation rates at least twice as high as households with above average net wealth holdings in Austria, Spain, and Italy - countries with low stock market participation rates. This revelation indicates that a more newly exploited angle of incidence should be focused upon, namely differences in country-specific factors and their influence on stock market participation.

Hong et al. (2004) report that social interaction is an important predictor for stock market participation and hence sociability matters for stock market participation. Brown, Ivković, Smith, and Weisbenner (2008) show a positive relation between community stock market participation and an individual's decision to participate in the stock market. Since a sociable community affects the stock market participation decision, financial literacy could turn out to be less important in the participation decision. A country-specific sociability factor which causes financial literacy to be less influential in predicting stock market participation indicates that the country-specific factor evidently moderates the existing relation. For this reason, the main question comprises the effect of country-specific social connectedness on the well-established positive relation between financial literacy and stock market participation. This results in the following research question:

How is the relation between financial literacy and stock market participation affected by country-specific levels of social connectedness?

The data used for this research is obtained from the Survey of Health, Ageing and Retirement in Europe (SHARE) database, Wave 6 and Wave 5. The relevance of using combined data will be explained in Section 3.2. The results of this paper confirm the relation between financial literacy and stock market participation as being positive and significant. Furthermore, the results provide evidence for a negative and significant moderating effect of country-specific social connectedness. This effect indicates that the predicting power of

financial literacy on stock market participation can be partly substituted by country-specific social connectedness.

The remainder of this research is structured as follows. The literature review emphasizes the synthesis in existing literature and eventually creates empirical predictions forming the basis for drawing up hypotheses. The methodology section defines the variables and the research method used followed up by the data section describing the data requirements, descriptive statistics, and data validity. The penultimate section elaborates on the results derived from univariate and multivariate statistics supplemented with robustness tests. The final section provides the conclusion and discussion.

2. Literature review

2.1 Stock market participation and financial literacy

For decades, academics have been trying to get a better understanding of stock market participation and the parameters influencing individuals' decisions whether or not to participate in the stock market. More than 30 years ago, Mehra and Prescott (1985) researched equity investments and concluded that equity premiums are higher than what would be expected compared to risk-free government bonds. Nonetheless, multiple academics mentioned and showed relative low stock market participation rates in the years following (Mankiw and Zeldes, 1991; Haliassos and Bertaut, 1995). This phenomenon is better known as the stock market participation puzzle. Van Rooij et al. (2011) report that financial literacy influences financial decision making and hence individuals with a low financial literacy are less likely to participate in the stock market. Their conclusion causes concern since individuals nowadays increasingly have to rely on themselves regarding important financial decisions and the financial literacy of young adults is worryingly low (Lusardi, Mitchell, and Curto, 2010).

Financial literacy is decisive for creating wealth and Van Rooij, Lusardi, and Alessie (2012) consider this feasible via stock market participation. The authors complement their argumentation by indicating that financial literate individuals face lower costs for collecting and processing information and thus face a lower economic threshold for stock market participation. Campbell (2006) reports that non-participation and under-diversification, seen as investment mistakes, are dependent on the level of wealth and education. Moreover, he suggests that individuals aware of their limited investment skills are less likely to have any funds in the

equity market. This is in line with Graham, Harvey, and Huang (2009), whom indicate that individuals aware of their investment skills and feeling knowledgeable are more likely to participate in the stock market. This reasoning is in accordance with the findings of Guiso and Jappelli (2005) whom indicate that a lack of financial awareness contributes to the explanation of the relatively low participation rates in financial markets. The authors conclude that stock market participation rates would increase considerably if all investors were fully aware of risky assets. The findings together underline that a lack of cognitive abilities contributes to the explanation of confined stock market participation rates, which is shown by Christelis, Jappelli, and Padula (2010). The authors state that an increase in an individual's cognitive ability lowers the information cost for both direct and indirect stock market participation. Additionally, they find this relation to be weaker compared to bondholding suggesting that the relation between cognitive abilities and stock market participation is most likely driven by informational constraints. In other words, the decision to participate in the stock market is linked to the ease or cost at which the information required is available to a certain individual.

Academics seem to agree on the fact that an individual's financial knowledge is important in predicting stock market participation. Education in finance helps improving financial literacy (Baker and Ricciardi, 2014) and some even claim that motivation is a key factor in determining financial literacy (Mandell and Klein, 2007). Furthermore, Jappelli and Padula (2013) emphasize that improving numeracy early in life eventually increments wealth accumulation. The importance of financial literacy is further established in relation to financial turmoil since individuals with a greater financial literacy are less likely to suffer damage from unexpected macroeconomic shocks (Klapper, Lusardi, and Panos, 2013). Taken together, financial literacy contributes to the explanation of the stock market participation puzzle even though major cross-country differences remain (Jappelli, 2010). The proposition of financial literacy being positively related to stock market participation leads to the following hypothesis:

H_1: Financial literacy positively influences stock market participation.

2.2 Stock market participation and country-specific social connectedness

Aside from the generally low level of stock market participation, the stock market participation puzzle also relates to large cross-country differences. Although stock market participation rates have increased over time, substantial differences in participation rates across countries remain (Guiso et al., 2003). Similar findings are acquired by Georgarakos and Pasini (2011) whom indicate that stock market participation rates differ among individuals in Europe

despite the comparable wealth holdings across countries. The authors recommend policy makers to encourage improvement of financial literacy and to improve transparency in financial markets which eventually promotes stockholdings in countries where stock market participation rates are rather low. The authors acknowledge the importance of regional disparities and assume that factors broader than an individual level could affect stock market participation rates as well. This reasoning is consistent with Thomas and Spataro (2015) whom, after observing heterogeneous capital market participation within and between countries, conclude that country-level institutional variables contribute to the explanation of the still rather limited participation puzzle. Stock market participation is mainly related to the individual perspective, however it seems that more broadly oriented factors account for the stock market participation puzzle as well.

As mentioned before, Christelis et al. (2010) indicate that stocks are information intensive assets and stockholding decisions depend on the accessibility or limitations of information. The informational constraint can be embodied both via the absence of information and the disability to process information. Since stock market participation is dependent on information available to an individual, having a lower financial literacy could be partly offset by information gathered through other channels. Accordingly, socio-economic factors could contribute to the explanation of deviations in stock market participation rates per country as well. Hong et al. (2004) hypothesize that social interaction contributes to an increase in stock market participation. They test the effect of sociability by either interacting with neighbors or attending church and find a positive relation with stock market participation. The authors propose both a decrease in information costs due to information sharing by word-of-mouth and the enjoyment people experience by talking about stock market investments as possible explanations for stock market participation. The first argument corresponds with the findings of Georgarakos and Pasini (2011) whom conclude that sociability decreases the cost of information by means of information sharing. The second argument, talking together about the market, could incentivize people to start acting like their peers (Duflo and Saez, 2002; Kaustia and Knüpfer, 2012). Duflo and Saez (2002) state that, with reference to retirement plans, the financial decisions individuals make are heavily dependent on peer effects. The authors show that decisions taken by someone within a peer group influences the decisions of another person in the same peer group due to social norms and a learning effect. Despite recognition of the importance, the distinction between social norms or a social learning effect is not made in their research. Kaustia and Knüpfer (2012), whom argue that peer performance does affect stock

market participation, declare that in countries or regions where social learning takes place the effect of peer performance is even stronger. The arguments are thorough to such extent that entry of peers is not affected by realized returns falling to zero or even below zero since people are less willing to talk about their inferior results compared to their investment victories. Han and Hirshleifer (2016) refer to this phenomenon as a self-enhancing transmission bias. Moreover, by taking into account the social and peer effects, simply having a larger social network increases the probability of having peers within that particular network that are participating in the stock market. Thus, the probability of an individual participating in the stock market will increase with the size of a social network.

Brown et al. (2008) state that community effects matter for stock market participation and participation increases when a larger fraction of an individual's community holds stocks. This effect is pronounced in more sociable communities caused by the word-of-mouth communication. Similar results are found in studies such as Hong, Kubik, and Stein (2005), and Christelis et al. (2010), albeit in the latter named an information spillover. Changwony, Campbell, and Tabner (2015) suggest that intensity and variety of social interactions both influence the level of stock market participation. More specific, the authors state that interactions via weak ties, which are seen as connections with formal and informal organizations, are more effective determinants for stock market participation than interactions via strong ties, which are seen as connections via neighbors, family, and close associates. These 'ties' relate to Granovetter's theory of social networks. Granovetter (2005) explicates that social networks normally have more weak ties than strong ties and weak ties are essential in "transmitting unique and non-redundant information across otherwise largely disconnected segments of social networks" (p. 35). Sociability positively influences stock market participation, however generalizing these results is precarious since part of the literature seems to indicate that stock market participation is influenced by strong ties as peers, colleagues, and family, whereas others indicate that rather weak ties are deterministic in stock market participation. Furthermore, as one can imagine, countries are different and a country-specific institutional context shapes the social networks accordingly (Kiss and Danis, 2008). The authors conducted research in entrepreneurial social networks and they conclude that strong ties are more prevailing at lower levels of institutional development whereas weak ties are more prevailing at higher levels of institutional development.

Alan (2006) shows that the entry costs of stock market participation support the explanation of the stock market participation rates and Guiso et al. (2003) show country

differences in stock market participation rates to be dependent on the existence of participation costs. The authors base these deviations across countries on institutional differences and informational barriers. All in all, the relation between financial literacy and stock market participation, which is based on the knowledge of an individual, should allow for some institutional differences. Financial literacy seems less relevant in countries where information required to participate in the stock market can be obtained more easily via an individual's surroundings. Since the variety and intensity of all relationships seem to be relevant for stock market participation, the prevailing level of social connectedness as macroeconomic factor could partly substitute the need of individual financial literacy. Hence, it is hypothesized that financial literacy becomes less decisive in explaining the stock market participation in countries with higher social connectedness. This proposition leads to the following hypothesis:

H_2: Country-specific social connectedness moderates the relation between financial literacy and stock market participation.

3. Methodology

3.1 Database

The data required for this research will be obtained from the Survey of Health, Ageing and Retirement in Europe (SHARE). SHARE[1] is a multidisciplinary and cross-national panel database containing micro-level data. The data originated from interviewing households consisting of a main respondent aged 50 years or older having its residence in Europe. The interviews relate to topics as health, socio-economic status, and social networks. The DNB Household Survey (DHS) provides similar financial data on 2000 households, however, the DHS is solely based on Dutch households. Since this research aims to provide insights on a multinational scale, SHARE data is most suitable.

The respondents of SHARE are mainly aged 50 years and over, which causes a skewed reflection of the population in a country. Although this could be seen as a constraint, Georgarakos and Pasini (2011) argue that households consisting of older people control a substantial part of society's resources. Consequently their degree of investments has substantial economic meaning within the macroeconomic spectrum. The SHARE database is continuously

[1] See References: Börsch-Supan (2017); Börsch-Supan, Brandt, Hunkler, Kneip, Korbmacher, Malter, Schaan, Stuck, Zuber (2013).

expanding the number of respondents and countries involved. Furthermore, new questions and modules are added resulting in broader and more extensive implications. Using the most recent data is pertinent since changes in thoughts or actions could have occurred in or after the financial crisis in 2007 and the euro crisis in 2009. Moreover, doing research with recent data is better grounded to contemporary phenomena.

The data collection for SHARE Wave 6 was completed in November 2015 and the data set was released on March 31, 2017. The respondents incorporated in SHARE Wave 6 reside in 18 countries, denominated: Austria, Belgium, Croatia, Czech Republic, Denmark, Estonia, France, Germany, Greece, Israel, Italy, Luxembourg, Poland, Portugal, Spain, Sweden, Switzerland, and Slovenia.

3.2 Variable constructions

3.2.1 Stock market participation

The dependent variable in this research is stock market participation. The SHARE database contains multiple questions relating to investments in riskless and risky assets. SHARE defines stock market participation as some form of investment which allows a person to own part of a corporation and gives a person the right to receive dividends from it. Stock market participation is divided into direct and indirect stock market participation. In line with Georgarakos and Pasini (2011), and Christelis et al. (2010), direct stock participation is derived from the following question stated in the questionnaire:

1.1 "Do you currently have any money in stocks or shares that are listed or unlisted on the stock market?"

The answer can be either *"yes"* or *"no"*. If the respondent answers yes, the respondent is considered a direct stock market participant. An individual could also participate in the stock market by means of a mutual fund. Mutual funds could consist of any type of assets, nonetheless in the questionnaire is chosen for a rather general distinction, namely stocks and bonds. The following question is used in determining whether a respondent participates in a mutual fund:

1.2 "Do you currently have any money in mutual funds or managed investment accounts?"

The answer can be either *"yes"* or *"no"*. If the respondent answers no, the answer is defined as no indirect stock market participation. If the respondent answers yes it is conceivable that the participant engages in a mutual fund which mainly, or even solely, invests in bonds. Christelis

et al. (2010) argue that stocks are more information intensive securities than bonds and the relation between financial literacy and stock market participation is most likely driven by the informational constraints. Hence, the distinction between stocks and bonds is important. Despite recognizing its importance, Christelis et al. (2010) assume that any mutual fund contains some stock and thereby do not specify the composition of a mutual fund. However, the authors state that the SHARE database does not give insight into the amount of stocks held via mutual funds but since this is possible nowadays the distinction is taken into account. Therefore, in this research, the precondition for taking a mutual fund into account is a mutual fund to comprise a reasonable degree of stocks. The questionnaire contains a question disclosing the composition of the mutual fund an individual is investing in. If the respondent answered yes on the previous question (*1.2*), the question is followed up by the following question:

1.3 "Are these mutual funds and managed investment accounts mostly stocks or mostly bonds?"

The respondent has three options: "*mostly stocks; half stocks and half bonds; mostly bonds"*. If at least half of the mutual fund consists of stocks, an individual engages in the stock market and faces substantial exposure on the equity market. Put differently, an individual answering mostly stocks or half stocks and half bonds is defined as indirect stock market participant. Contrary, if the respondent answers mostly bonds, the answer is defined as no indirect stock market participation. Finally, indirect stockholdings via employee or retirement programs are not included in this research because individual investors do not control these investments.

The total stockholdings, the direct and indirect stock market participation together, is taken into consideration for determining the dependent variable in this research. If an individual engages in direct stock market participation, engages in indirect stock market participation or engages in both, the respondent will be marked as stock market participant. All other respondents are considered non-stock market participants and thus the dependent variable is a dichotomous variable.

3.2.2 Financial literacy

Financial literacy is an explanatory variable in this research. There is little ambiguity in current literature that a proper proxy for financial literacy should entail some calculations. Although SHARE is not as extensive on this construct as for instance the DHS, the SHARE

questionnaire contains some financially framed numeracy questions. Wave 6 contains limited data on numeracy since numeracy questions were solely asked to baseline respondents, that is, respondents whom participate in SHARE for the first time. Baseline respondents result from an extension in or within countries. The longitudinal respondents, those who participated and answered questions relating to numeracy in an earlier wave, do not have to answer the questions again. Therefore, Wave 6 contains a large number of missing values. SHARE does not consider the questions relating to numeracy subject to change and therefore this research complements data from Wave 6 with data from Wave 5 if necessary. SHARE Wave 5 contains abundant data on numeracy and the collection of data was completed in November 2013.

Numeracy is associated with financial literacy and both affect an individual's financial decision-making (Christelis et al., 2010; Van Rooij et al., 2011). Moreover, numerical skills are a predictive proxy for wealth which ultimately is partly held in stocks (McArdle, Smith, and Willis, 2009). Lusardi and Mitchell (2007) assess financial literacy via a percentage question, a division question, and an interest compounding question. In the SHARE questionnaire, only the division question is slightly different compared to what is used by the authors. Furthermore, the SHARE questionnaire contains one additional division question, but since the remaining questions are equivalent to what the authors use for the determination of financial literacy, the following questions are used for this research:

2.1 "If the chance of getting a disease is 10 percent, how many people out of 1,000 would be expected to get the disease?"

2.2 "In a sale, a shop is selling all items at half price. Before the sale, a sofa costs 300. How much will it cost in the sale?"

2.3 "A second-hand car dealer is selling a car for 6,000. This is two-thirds of what it costs new. How much did the car cost new?"

2.4 "Let's say you have 2,000 in a savings account. The account earns ten per cent interest each year. How much would you have in the account at the end of two years?"

A respondent's level of financial literacy will be determined by the number of correct numeracy questions. Dewey and Prince (2005) constructed the design of this measure. The first question (*2.1*) is at intermediate level and the next question depends on whether the respondent is able to answer the first question correct. If the respondent answers the first question incorrect, the next question will be the second question (*2.2*). This question is considered easiest and if the respondent answers this question incorrect he or she receives the score of 1 on financial literacy.

The respondent is then considered financially illiterate. If the respondent answers the second question (*2.*2) correct, he or she receives a score of 2 on financial literacy and the questions will stop at this point for the respondent. When the respondent answers the first question (*2.1*) correct, the respondent already accomplished the intermediate level indicating that he or she receives a score of 3. The respondent may try the third question (*2.3*) at an advanced level and if the respondent answers this question incorrect, the financial literacy score will remain 3. Contrary, if the respondent answers the third question (*2.3*) correct, the financial literacy score increases to 4 and the respondent may try the last and most difficult question (*2.4*) which relates to compounding interest. If the respondent answers the last question (*2.4*) incorrect the financial literacy score will remain 4 and if the respondent answers the last question correct, he or she will receive a score of 5 which indicates a high financial literacy. With an increase or decrease of financial literacy is meant an increase or decrease in the financial literacy score, henceforth financial literacy and financial literacy score could be used interchangeably.

3.2.3 Social connectedness

The moderating effect of country-specific social connectedness on the relation between financial literacy and stock market participation will be elaborated upon in this research. To explain this moderating effect on stock market participation, an interaction term will be added to the model. Social connectedness is a newly introduced summary scale on social network data in Wave 6. Data on social connectedness[2] is not the result of a specific question asked in the questionnaire, but a generated variable calculated from five components. More specifically, it considers a social network's size, diversity, proximity, contact, and emotional closeness. Social connectedness is relatively new and scarce in the existing literature making the comparison between different measures challenging. However, in accordance with Bailey, Cao, Kuchler, Stroebel, and Wong (2017), size, diversity, and proximity are important determinants for social connectedness.

The components of social connectedness will be discussed using the initial questions from the questionnaire. Firstly, a respondent is asked to list important people to him or her by mentioning their first name. This component provides insight into the size of the respondent's social network and the score is a sum of the number of members which could vary between 0 and 7. Subsequently, for the network diversity component the respondent specifies its

[2] I would like to thank Ella Schwartz and Kim Stoeckel, authors of the 'social connectedness' syntax, for sharing the syntax, which enabled assessment of suitability of the indicator for this research.

relationship by appointing each of the aforementioned social network members into a category that reads: spouse, other family, friend, or all others. The score of the diversity component is the sum of the number of different categories mentioned by the respondent and could vary between 0 and 4. The following component relates to the geographic proximity of each of the social network members. A network member is defined as in close proximity if he or she lives within 25 kilometers of the respondent. The score of a respondent's geographic proximity is defined by determining how many members live within 25 km of the respondent and could vary between 0 and 7. Hereafter, the component network contact should give insight in how often a respondent contacts each separate member in person, by phone, by email, or by any other electronic means. Weekly contact or even more is defined as intensive contact. The score of network contact is defined by determining with how many members the respondent has weekly contact or more and this could vary from 0 to 7. Finally, a respondent is asked what its emotional closeness towards each particular network member is. The emotional closeness is considered high if the respondent answers to be very or extremely close towards a social network member. The score of emotional closeness is defined by the number of members which are very or extremely close to the respondent and could vary between 0 and 7.

Prior to generating the social connectedness variable, the five components must be computed separately. For size, proximity, amount of contact, and emotional closeness, 0 social network members generates the score 0; 1 social network member generates the score 1; 2 or 3 social network members generate the score 2; 4 or 5 social network members generate the score 3; and 6 or 7 social network members generate the score 4. The diversity component is calculated in a divergent manner and the score is composed as the sum of the dummies attached to the categories. For example, if the respondent has 4 social network members and they are all friends, the diversity score is 1. If the respondent has 4 social network members whereof 3 are friends and 1 is its spouse, the diversity score is 2. Therefore, a higher score on the diversity component suggests more diversity in a respondent's social network. To generate the social connectedness score, all five separate components are summed which results in a value between 0 and 20. Hence, a higher score represents a higher social connectedness.

Until here, the generated social connectedness score is on individual grounds, however this research aims at country-specific differences. Brown et al. (2008) argue that defining the geographic size of a variable is challenging since it must be sufficiently large to capture all social groups and human interconnections, but not excessively large that all potential observed effects will be diluted. Furthermore, the authors declare that most micro-level databases do not

fit this motive perfectly. However, as noted in Section 2.2, cross-country differences seem relevant in the explanation of stock market participation and therefore the effect of country-specific social connectedness is measured via the creation of a country-average. Utilizing a country-level measure is in accordance with Thomas and Spataro (2015) although they use the effectiveness of a country's education system and the attractiveness of a country's financial market to explain differences in stock market participation rates. Since country-averages could mutually diminish pronounced effects within a country, a robustness check should reveal whether the results are maintained if the variable social connectedness is replaced by the variable social cohesion. The latter variable is derived from a rather macroeconomic source and a more detailed explanation follows in Section 5.3.

3.2.4 Control variables

To accurately measure the moderating effect of country-specific social connectedness on the relation between financial literacy and stock market participation, controlling for additional variables affecting the variance of stock market participation is a prerequisite. In line with multiple scholars in this field (see, e.g., Van Rooij et al., 2011; Georgarakos and Pasini, 2011) there will be controlled for age. For the univariate statistics, age will be subdivided into three groups. These groups are roughly divided in the working-age population (<59), the young retirees (60 to 79), and the elderly population (>80). For the multivariate statistics, which are used for the ultimate regression model, age in years is used since this measure entails more predictive power and does not result in any loss of data. Furthermore, retirement is examined to be a control variable, nonetheless, age and retirement 'compete' to explain the same variance in the ultimate regression model and the high correlation between age and retirement is decisive in excluding retirement as control variable.

A recent study confirmed that gender still influences the likelihood of stock market participation (Almenberg and Dreber, 2015). Additionally, Van Rooij et al. (2011), and Georgarakos and Pasini (2011) control for gender. Gender is in both univariate and multivariate statistics a dummy variable, indicating a 1 for male respondents and a 0 for female respondents.

Rosen and Wu (2004) find evidence for holding a larger part of wealth in safe assets instead of risky assets caused by an individual's poor health condition. Accordingly, controlling for health is relevant in the context of stock market participation. The SHARE database provides numerous proxies for health, however in line with Rosen and Wu (2004), the health variable is based on an individual's self-reported health. The self-reported health score varies

from 1 to 5 indicating an excellent and poor health respectively. Since a higher score corresponds to a poorer health, a negative relation is expected. The reasoning of using self-reported health is strengthened by the overly general manner of measuring long-term illness in the SHARE questionnaire. The question relates to having any chronic health problems, illnesses, or disabilities. However, an individual could have bronchitis or could use a wheelchair which is likely to affect the individual in its daily activities although these conditions do not necessarily have to affect stock market participation. Other academics use depression as explanatory variable. However, the SHARE questionnaire asks the respondent if he or she was sad or depressed in the last month and through this question, no less than 38.5% of all respondents were considered depressed. Hence, the measure depression turns out to be inferior to self-reported general health.

The majority of the respondents is about retirement age, a substantial part of the respondents is not receiving income through work and if they receive income it could be via interest, pension plans, or government payments. To avoid ambiguity about what to define as income and what not, the more comprehensive measure wealth is used. In line with the arguments of Hong et al. (2004), both wealth and education are substantial predictors of stock market participation and thus should be controlled for. Wave 6 contains data from 18 countries and using averages is crucial to circumvent difficulties in comparisons with different currencies. Moreover, due to purchasing power differences and currency fluctuations, absolute wealth is highly subjective when comparing multiple countries. For the univariate statistics, a country-average will be created and wealth will be grouped into above country-average and below country-average. For the multivariate statistics, which are used in the ultimate regression model, wealth is also relative, however to increase the explanatory power, an individual's wealth is determined by the creation of a fraction. By means of a fraction, the value 3 would indicate that the specific respondent is 3 times as wealthy as the average respondent in that particular country.

Scholars do agree on education influencing stock market participation rates as well (see, e.g., Hong et al., 2004; and Van Rooij et al., 2011). Education is measured in years and varies from 0 to 25. For education, the same constraint is encountered as for financial literacy. Those respondents who participated and answered questions relating to education in Wave 5 do not receive the questions again in Wave 6 and thus the latter contains a high share of missing values. Since SHARE considers the variable education to be constant, Wave 6 data is complemented with Wave 5 data if available.

Finally, many scholars suggest that actively participating in the stock market is correlated with an individual's risk preference (see, e.g., Vissing-Jørgensen and Attanasio, 2003; Van Rooij et al., 2011). Individuals willing and prepared to take on larger risks are more likely to invest in stocks. Contrarily, individuals not willing to take on extra levels of risk are said to be risk averse. As for financial literacy and education, SHARE decided not to assume risk aversion subject to change and missing data for risk aversion in Wave 6 is complemented with data from Wave 5 wherever available. Risk aversion is measured by the amount of financial risk an individual is willing to take when saving money or making investments. The score varies from 1 to 4, being respectively: substantial financial risk expecting to earn substantial returns; above average financial risk expecting to earn above average returns; average financial risk expecting to earn average returns; and not willing to take any financial risk. Hence, a negative relation is expected.

3.3 Regression models

To ensure validity and reliability of the moderating effect of country-specific social connectedness on the relation between financial literacy and stock market participation, the equation belonging to the first hypothesis is regressed. The first hypothesis solely measures the effect of financial literacy (FLT) on stock market participation (SMP). Furthermore, the relation between the independent variable and the dependent variable must be positive to comply with the existing literature and in order to make any legitimate inferences regarding social connectedness and its moderating effect on the aforementioned relation. As noted in Section 3.2.4, the regression model controls for age (AGE), gender (GEN), health (HEA), wealth (WEA), education (EDU), and risk aversion (RIA). The regression will be concluded with an error term (ε) and the individual respondent is represented by subscript ($_i$). Eq. (1) shows the model for testing the first hypothesis:

$$SMP_i = \alpha + \beta_1 FLT_i + \beta_2 AGE_i + \beta_3 GEN_i + \beta_4 HEA_i + \beta_5 WEA_i + \beta_6 EDU_i + \beta_7 RIA_i + \varepsilon_i \quad (1)$$

For the second hypothesis the model contains, besides an individual's financial literacy (FLT), country-specific social connectedness (SCN) itself and an interaction term (FLT*SCN). The interaction term consists of financial literacy and country-specific social connectedness and represents the moderating effect of the second hypothesis. Similar to the first hypothesis, the regression model controls for age (AGE), gender (GEN), health (HEA), wealth (WEA), education (EDU), and risk aversion (RIA). Again, the regression will be concluded with an

error term (ε). The individual respondent is represented by subscript ($_i$) where its country of residence is represented by subscript ($_c$). Eq. (2) shows the complete model and is used for testing the second hypothesis:

$$SMP_{ic} = \alpha + \beta_1 FLT_i + \beta_2 SCN_c + \beta_3 (FLT_i * SCN_c) + \beta_4 AGE_i + \beta_5 GEN_i + \beta_6 HEA_i + \beta_7 WEA_i + \beta_8 EDU_i + \beta_9 RIA_i + \varepsilon_{ic} \quad (2)$$

As mentioned before, stock market participation (SMP) is a dichotomous variable. A binary logistic regression produces values that predict the dependent variable based on the independent variables. However, the coefficients resulting from the binary logistic regression cannot be interpreted similarly as Ordinary Least Square (OLS) coefficients since they are generated in log-odds units. Therefore, to enhance the interpretation of the effects, the odds-ratio will also be presented. The odds-ratio, obtained via the exponentiation of the initial coefficient, is the change in odds of the dependent variable resulting from a 1-unit increase of an explanatory variable. Since a binary logistic regression is not linear and the probability of participating in the stock market is not constant at any marginal change, the odds-ratio provides the change in odds of stock market participation by a marginal change. To avoid any ambiguity, it is worth mentioning that the odds-ratio is not equivalent to the probability.

4. Data

4.1 Construction research sample

To obtain a comprehensive understanding of the first and the second hypothesis, the full sample consists of individuals participating via direct means, indirect means, or via direct and indirect means. This full sample is referred to as the total participation sample. Additionally, to provide some extra insights a subgroup is drafted with the sole requirement of direct stock market participation. An overview is presented in Table 1.

Table 1
Sample summary. This table presents an overview of the initial number of respondents included in the SHARE database and the ultimate number of respondents taken into account for this research. Total participation consists of direct and/or indirect stock market participation and direct participation solely requires information on direct stock market participation.

	Total participation	**Direct participation**
Number of observations total database	68,231	68,231
(Missing values stock market participation)	(22,733)	(22,773)
Valid values stock market participation	45,498	45,458
(Missing values financial literacy)	(28,110)	(28,087)
(Missing values country-average social connectedness)	(0)	(0)
(Missing values control variables)	(470)	(470)
Number of respondents taken into account	**16,918**	**16,901**

A reliable regression model should entail complete data on the dependent variable, the independent variables, the interaction term and the control variables. The minor difference between the subgroup and the total participation sample is caused by some individuals having a missing value on direct stock market participation. In the total participation sample, additional data is available from indirect stock market participation and therefore the extra information can be included. The grouping variable in this research is country residence and every country must have sufficient valid observations to represent a country's specific social connectedness. Approximately 20 years ago, Peduzzi, Concato, Kemper, Holford, and Feinstein (1996) did research on the minimum number of observations required for reliable logistic regression analysis. They concluded that less than 10 observations per variable could result in biased regression coefficients. More recent research conducted by Vittinghoff and McCulloch (2007) argued that rules of thumb with 10 observations per variables are signals for potential shortfalls. Moreover, the authors argue that for predictions, in this research stock market participation, a minimum of 20 observations may be more appropriate. Even more recently, Courvoisier, Combescure, Agoritsas, Gayet-Ageron, and Perneger, (2011) stipulate that no single rule of thumb could fully guarantee the estimations to be correct. Nonetheless, the authors advice that increasing the number of observations as well as being attentive with regards to the correlations between variables will benefit the results.

The subgroup of direct stock market participants consists of 16,901 complete observations and is of sufficient size to draw conclusions from. For another subgroup with the sole requirement of indirect stock market participants, running a reliable regression would be feasible only for 1,840 individuals spread over 18 countries. 6 countries would have 20 observations or less of which 4 countries would have 5 observations or less. Hence, based on the insufficient number of observations an extra subgroup with the sole requirement of indirect

stock market participation is excluded from this research. Besides the sample size, the sample representativeness is important as well and this will be discussed in Section 4.3

4.2 Descriptive statistics

The descriptive statistics of the total participation sample are presented in Table 2. The table contains country-specific means, standard deviations, minimums, maximums, and medians on all variables used for this research. When considering the direct participation subgroup, minor deviations in descriptive statistics would occur for the financial literacy variable and the control variables. However, the key contribution of this research covers the country-specific social connectedness and its interaction with financial literacy and since these values are not changing and not present respectively, solely the descriptive statistics for the full sample are provided. Moreover, the most pronounced differences are expected in the interaction term of financial literacy and country-specific social connectedness and hence comparing both samples in the multivariate results is more informative.

Table 2

Descriptive statistics. This table provides an overview of descriptive statistics for the main variables and the control variables of the total participation sample. Numbers are based on Wave 6 and Wave 5 of the SHARE database. Number of observations is 16,918. Standard deviations are provided in parentheses. For complete variable overview see Appendix A.1.

Country	N		SMP	FLT	SCN	AGE	GEN	HEA	WEA	EDU	RIA
		Mean (S.D.)	0.17 (0.38)	3.83 (1.15)	11.31 (0)	69.13 (10.20)	0.53 (0.50)	2.96 (1.10)	0.80 (0.95)	8.30 (4.37)	3.72 (0.50)
Austria	47	Min./Max.	0/1	1/5	11.31	51/91	0/1	1/5	-0.10/4.24	2/17	2/4
		Median	0	4	11.31	68	1	3	0.46	8	4
		Mean (S.D.)	0.18 (0.39)	3.66 (0.99)	10.91 (0)	65.33 (9.99)	0.47 (0.50)	3.19 (0.98)	0.98 (1.57)	12.71 (3.80)	3.69 (0.54)
Germany	2,160	Min./Max.	0/1	1/5	10.91	42/95	0/1	1/5	-0.78/18.39	0/25	1/4
		Median	0	4	10.91	64	0	3	0.54	12	4
		Mean (S.D.)	0.54 (0.50)	3.72 (0.96)	10.65 (0)	68.44 (9.99)	0.50 (0.50)	2.68 (1.11)	0.99 (1.18)	11.80 (4.06)	3.28 (0.76)
Sweden	1,384	Min./Max.	0/1	1/5	10.65	43/94	0/1	1/5	-0.39/10.83	0/25	1/4
		Median	1	4	10.65	68	0	3	0.66	12	3
		Mean (S.D.)	0.04 (0.19)	2.68 (1.02)	9.40 (0)	69.33 (11.26)	0.43 (0.50)	3.21 (1.01)	1.03 (1.21)	9.86 (5.14)	3.88 (0.40)
Spain	1,516	Min./Max.	0/1	1/5	9.40	36/100	0/1	1/5	-2.42/16.78	0/25	1/4
		Median	0	3	9.40	68	0	3	0.79	9	4
		Mean (S.D.)	0.05 (0.22)	3.12 (1.07)	8.72 (0)	63.70 (10.05)	0.43 (0.50)	3.10 (1.04)	0.94 (0.92)	9.89 (4.71)	3.67 (0.67)
Italy	1,524	Min./Max.	0/1	1/5	8.72	40/102	0/1	1/5	-0.18/7.13	0/25	1/4
		Median	0	3	8.72	62	0	3	0.75	8	4
		Mean (S.D.)	0.18 (0.39)	3.26 (1.06)	10.44 (0)	57.18 (9.25)	0.50 (0.50)	2.95 (1.06)	0.80 (0.81)	12.91 (4.50)	3.56 (0.63)
France	284	Min./Max.	0/1	1/5	10.44	36/95	0/1	1/5	-0.08/5.85	0/23	1/4
		Median	0	3	10.44	53	0	3	0.65	12	4
		Mean (S.D.)	0.37 (0.48)	3.80 (1.04)	10.82 (0)	63.51 (10.31)	0.50 (0.50)	2.42 (1.15)	1.00 (1.44)	13.40 (3.77)	3.27 (0.79)
Denmark	1,185	Min./Max.	0/1	1/5	10.82	42/97	0/1	1/5	-1.02/15.37	1/25	1/4
		Median	0	4	10.82	62	1	2	0.60	13	3
		Mean (S.D.)	0.02 (0.16)	3.36 (1.02)	8.90 (0)	64.32 (10.51)	0.45 (0.50)	2.86 (1.10)	0.89 (1.06)	10.09 (4.58)	3.78 (0.59)
Greece	1,675	Min./Max.	0/1	1/5	8.90	41/95	0/1	1/5	-1.53/12.14	0/25	1/4
		Median	0	3	8.90	63	0	3	0.60	10	4
		Mean (S.D.)	0.17 (0.38)	3.86 (0.88)	10.74 (0)	67.14 (8.19)	0.59 (0.50)	2.48 (0.87)	.90 (1.19)	9.62 (5.30)	3.72 (0.45)
Switzerland	29	Min./Max.	0/1	2/5	10.74	51/84	0/1	1/4	0.01/4.73	2/21	3/4
		Median	0	4	10.74	67	1	3	0.49	10	4
		Mean (S.D.)	0.25 (0.43)	3.55 (1.04)	10.68 (0)	62.20 (10.90)	0.48 (0.50)	2.89 (1.00)	1.00 (1.27)	12.58 (4.12)	3.62 (0.64)
Belgium	1,420	Min./Max.	0/1	1/5	10.68	24/95	0/1	1/5	-0.30/19.90	0/25	1/4
		Median	0	4	10.68	59	0	3	0.71	12	4

Table 2
Descriptive statistics – continued.

Country	N		SMP	FLT	SCN	AGE	GEN	HEA	WEA	EDU	RIA
		Mean (S.D.)	0.12 (0.33)	3.80 (1.09)	9.75 (0)	61.86 (8.62)	0.49 (0.50)	2.69 (1.16)	0.99 (1.32)	13.61 (2.96)	3.68 (0.67)
Israel	167	Min./Max.	0/1	1/5	9.75	39/90	0/1	1/5	-0.16/9.43	6/21	1/4
		Median	0	4	9.75	60	0	3	0.54	14	4
		Mean (S.D.)	0.09 (0.29)	3.67 (1.00)	9.35 (0)	66.64 (9.43)	0.33 (0.47)	3.28 (1.02)	1.03 (1.20)	12.28 (3.37)	3.55 (0.65)
Czech Republic	676	Min./Max.	0/1	1/5	9.35	39/103	0/1	1/5	-0.25/9.30	1/23	1/4
		Median	0	4	9.35	67	0	3	0.72	12	4
		Mean (S.D.)	0.02 (0.16)	3.33 (1.10)	8.88 (0)	54.76 (2.73)	0.43 (0.50)	3.29 (1.03)	1.31 (1.40)	12.02 (2.81)	3.76 (0.55)
Poland	245	Min./Max.	0/1	1/5	8.88	48/67	0/1	1/5	-1.12/8.62	1/22	1/4
		Median	0	3	8.88	55	0	3	0.92	12	4
		Mean (S.D.)	0.18 (0.38)	3.51 (1.07)	8.79 (0)	65.60 (9.64)	0.46 (0.50)	3.01 (1.02)	1.00 (1.05)	12.14 (4.44)	3.71 (0.57)
Luxembourg	1,129	Min./Max.	0/1	1/5	8.79	48/97	0/1	1/5	-0.10/17.51	0/25	1/4
		Median	0	4	8.79	64	0	3	0.77	12	4
		Mean (S.D.)	0.08 (0.27)	3.08 (1.05)	9.75 (0)	65.12 (9.24)	0.66 (0.48)	3.50 (0.97)	0.99 (1.18)	7.90 (5.23)	3.82 (0.56)
Portugal	50	Min./Max.	0/1	1/5	9.75	51/85	0/1	1/5	-0.64/4.63	3/25	1/4
		Median	0	3	9.75	61	1	4	0.55	6	4
		Mean (S.D.)	0.08 (0.27)	3.25 (0.97)	8.92 (0)	66.91 (9.91)	0.41 (0.49)	3.32 (1.03)	0.95 (0.96)	10.44 (3.47)	3.84 (0.40)
Slovenia	1,410	Min./Max.	0/1	1/5	8.92	47/95	0/1	1/5	-1.11/7.63	0/23	1/4
		Median	0	3	8.92	66	0	3	0.74	11	4
		Mean (S.D.)	0.04 (0.21)	3.41 (0.92)	9.29 (0)	54.05 (5.92)	0.38 (0.49)	3.37 (0.98)	1.31 (2.04)	12.61 (3.63)	3.62 (0.72)
Estonia	475	Min./Max.	0/1	1/5	9.29	39/89	0/1	1/5	-0.37/17.61	1/23	1/4
		Median	0	3	9.29	53	0	3	0.58	12	4
		Mean (S.D.)	0.04 (0.20)	3.25 (0.94)	10.14 (0)	65.47 (9.52)	0.40 (0.49)	3.33 (1.18)	0.97 (1.54)	10.23 (3.91)	3.77 (0.51)
Croatia	1,542	Min./Max.	0/1	1/5	10.14	29/95	0/1	1/5	-5.16/14.66	0/25	1/4
		Median	0	3	10.14	64	0	3	0.53	11	4
		Mean (S.D.)	0.16 (0.36)	3.40 (1.06)	9.79 (0.86)	64.90 (10.44)	0.45 (0.50)	3.04 (1.09)	0.99 (1.28)	11.40 (4.36)	3.66 (0.63)
Total	16,918	Min./Max.	0/1	1/5	8.72/11.31	24/103	0/1	1/5	-5.16/19.90	0/25	1/4
		Median	0	3	9.40	63	0	3	0.67	12	4

4.3 Research sample validity

Assuming that SHARE carefully compiled its respondents for all countries, the total participation sample should be a representative reflection of the total database. To provide insight in the validity of the sample, the means and standard deviations of the sample must be compared to the means and standard deviations of the total database. An overview is presented in Table 3.

Table 3
Sample validity. This table presents the means and standard deviations of the total participation sample and the total SHARE database. Numbers are based on Wave 6 and Wave 5 of the SHARE database. The research sample used consists of 16,918 respondents and the total database consists of 68,231 respondents. For every t-statistic, the p-value is provided in parentheses.

	Total participation sample		**Total database**		
	Mean	St. dev.	Mean	St. dev.	t-test (p-value)
Stock market participation	0.16	0.36	0.15	0.35	2.580 (0.010)
Financial literacy	3.40	1.06	3.37	1.07	2.598 (0.009)
Social connectedness	9.79	0.86	9.85	0.84	-6.492 (0.000)
Age	64.90	10.44	67.68	10.31	-24.644 (0.000)
Gender	0.45	0.50	0.44	0.50	1.839 (0.066)
Health	3.04	1.09	3.20	1.07	-13.625 (0.000)
Wealth	0.99	1.28	1.00	1.50	-0.660 (0.510)
Education	11.40	4.36	11.30	4.32	2.119 (0.034)
Risk aversion	3.66	0.63	3.69	0.59	-4.521 (0.000)
Number of observations	16,918		68,231		

At first sight, the means and standard deviations do not appear to deviate in a large extent. To statistically test this, an independent samples t-test is conducted. The null hypothesis indicates that the mean derived from the total database is equal to the mean of the research sample used. The smaller the p-value, the stronger the evidence against the null hypothesis. Except for wealth and gender, all null hypotheses are statistically rejected at the 5% significance level. Hence, from a statistical point of view, appropriate caution is required for the majority of the variables in generalizing the overall results of the research sample used.

5. Results

5.1 Univariate results on stock market participation

A high or perfect relation between predicting variables preclude estimates of the ultimate regression coefficients to be reliable and hence results in wrongful conclusions. This phenomenon is called multicollinearity. To exclude the presence of multicollinearity, a

bivariate correlation analysis is conducted. An overview of the results is provided in Table 4. The table shows coefficients which theoretically could range anywhere from -1 to 1, being a perfect negative relation or a perfect positive relation respectively. A coefficient of 0 would indicate no relation at all. A high correlation coefficient indicates multicollinearity and rules of thumb from basic statistical textbooks indicate that correlation coefficients ranging from 0.500 up till 0.700 are identified as highly correlated with 0.700 being the critical boundary.

Table 4

Correlation matrix. This table provides an overview of the correlations coefficients between sets of variables present in the model. ** and * indicate the 1% and 5% significance level respectively. Numbers are based on Wave 6 and Wave 5 of the SHARE database. Number of observations is 16,918. For complete variable overview see Appendix A.1.

Variable	FLT	SCN	AGE	GEN	HEA	WEA	EDU	RIA
FLT	1							
SCN	0.166**	1						
AGE	-0.220**	0.004	1					
GEN	0.181**	0.048**	0.011	1				
HEA	-0.239**	-0.085**	0.252**	-0.080**	1			
WEA	0.174**	0.001	-0.064**	0.087**	-0.149**	1		
EDU	0.405**	0.188**	-0.286**	0.104**	-0.262**	0.218**	1	
RIA	-0.210**	-0.158**	0.141**	-0.156**	0.185**	-0.173**	-0.224**	1

As mentioned earlier, the correlation coefficient of age and retirement is above 0.800 and hence retirement is excluded from the model. Coefficients being 0.300 up till 0.500 indicate moderate correlation and only the correlation coefficient of financial literacy and education is moderately correlated (0.405) and statistically significant at the 1% level. Coefficients being 0.100 up till 0.300 indicate weak correlation and coefficients being 0 up till 0.100 indicate very weak correlation or no correlation at all. Hence, all other sets of predicting variables show a weak or very weak correlation. Concluding, the absence of noteworthy correlation is confirmed and thus no changes in the existing model are required.

The probability of participating in the stock market depends on multiple variables. For univariate statistics, the starting point is ‘being a stock market participant’. Table 5 provides an overview containing different levels of financial literacy, country-average social connectedness, and six complementary socio-economic variables. Table 5 provides these ratios on the subgroup direct stock market participation as well. Chances of participating in the stock market increase with an increase in financial literacy. Fig. 2 is a visualization of the fraction stock market participation plotted against the financial literacy score.

Table 5

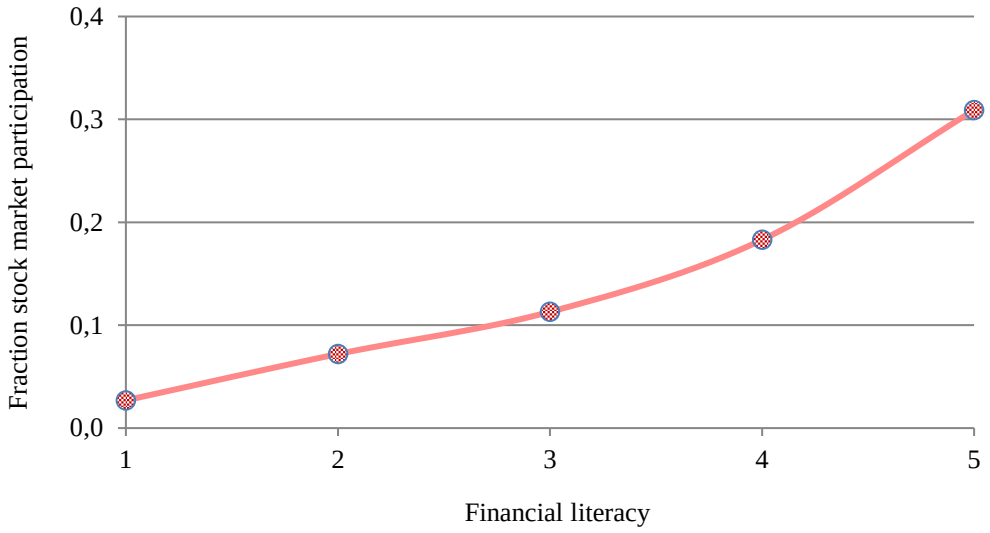
Univariate results. This table provides a univariate analysis on stock market participation on the basis of financial literacy, country-average social connectedness, and all control variables. Numbers are based on Wave 6 and Wave 5 of the SHARE database. Total SMP consists of 16,918 observations, whereof 2,649 participate and 14,269 do not participate in the stock market. Direct SMP consists of 16,901 observations, whereof 15,074 participate and 1,827 do not participate in the stock market. For every t-statistic, p-value is provided in parentheses.

	Total SMP	Direct SMP		Total SMP	Direct SMP
Financial literacy			**Health**		
1 (low)	2.7%	2.1%	1 (excellent)	27.6%	20.5%
2	7.2%	4.8%	2	22.2%	15.7%
3	11.3%	7.9%	3	15.7%	10.2%
4	18.3%	12.4%	4	9.5%	6.7%
5 (high)	30.9%	22.0%	5 (poor)	5.6%	3.5%
Δ high - low	28.2%	19.9%	Δ poor - excellent	-22.0%	-17.0%
t-test (p-value)	30.144 (0.000)	24.753 (0.000)	t-test (p-value)	-23.129 (0.000)	-20.677 (0.000)
Country-average social connectedness			**Wealth**		
Q1 (low)	7.3%	4.7%	Below average (low)	9.5%	6.2%
Q2	4.6%	3.6%	Above average (high)	27.9%	20.0%
Q3	26.8%	19.4%	Δ high - low	18.4%	13.8%
Q4 (high)	24.5%	16.8%	t-test (p-value)	23.687 (0.000)	20.611 (0.000)
Δ high - low	17.2%	12.1%	**Education**		
t-test (p-value)	38.834 (0.000)	33.485 (0.000)	Q1 (low)	4.2%	3.0%
Age			Q2	12.0%	8.0%
< 59 (young)	14.3%	10.0%	Q3	24.1%	16.8%
60 to 79	17.4%	11.9%	Q4 (high)	29.6%	21.4%
> 80 (old)	11.8%	8.3%	Δ high - low	25.4%	18.4%
Δ old - young	-2.5%	-1.7%	t-test (p-value)	28.184 (0.000)	23.721 (0.000)
t-test (p-value)	0.181 (0.857)	0.090 (0.929)	**Risk aversion**		
Gender			1 (low)	30.0%	21.9%
Male	20.4%	14.2%	2	47.7%	36.2%
Female	11.8%	8.1%	3	33.4%	24.0%
Δ male - female	8.6%	6.1%	4 (high)	8.3%	5.3%
t-test (p-value)	15.511 (0.000)	12.732 (0.000)	Δ high - low	-21.7%	-16.6%
			t-test (p-value)	-34.082 (0.000)	-29.951 (0.000)

Considering Table 5, the difference between the highest and the lowest financial literacy score is accompanied by a major increase of participating in the stock market for both the main research sample and the subgroup, namely 28.2% and 19.9% respectively. At first sight, a positive relation between financial literacy and stock market participation is observed supporting hypothesis 1. A figure showing the financial literacy scores per country is provided in Appendix A.2.

Figure 2

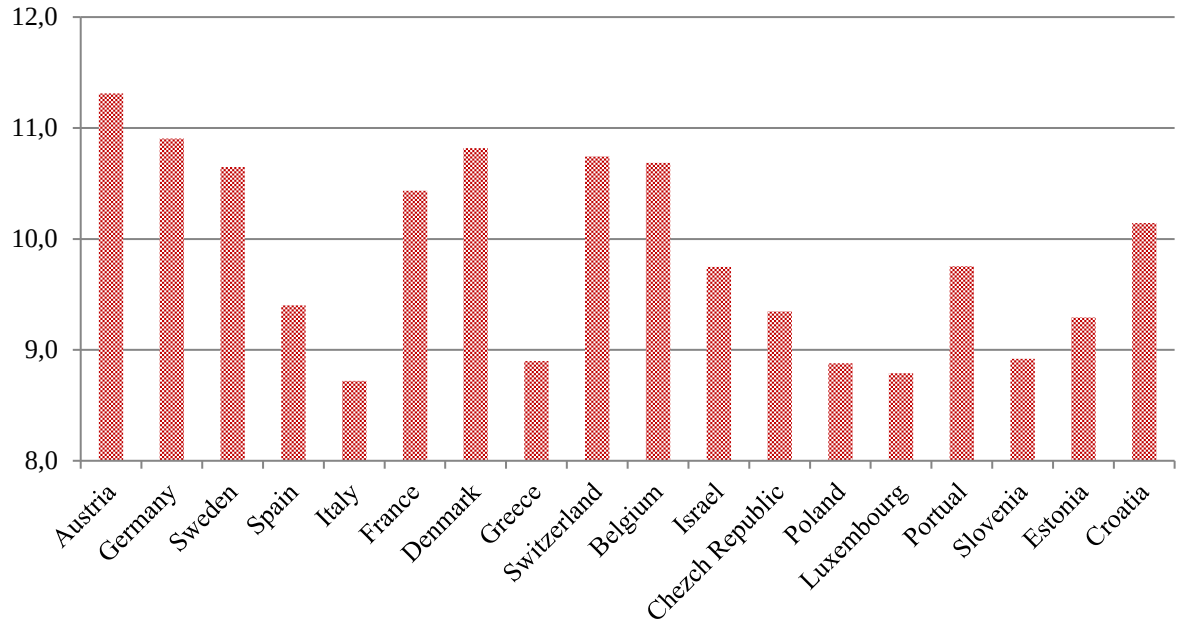
Financial literacy versus stock market participation. This figure provides the fraction of stock market participation compared to financial literacy scores. Numbers are based on Wave 6 and Wave 5 of the SHARE database and stock market participation contains both levels of direct and indirect stockholdings.



The country-average social connectedness is divided into four equal quarters with the lowest score being 8.72 (Italy) and the highest score being 11.31 (Austria). These values are consistent with the minimum and the maximum provided in the descriptive statistics, Table 2. At first sight, country-specific social connectedness shows a positive relation with stock market participation for both the main research sample and the subgroup. An overview of social connectedness per country is visualized in Fig. 3. The control variables in this research relate to six socio-economic variables and since these variables do not show noteworthy deviations in their relation towards stock market participation with reference to the current literature, these are explained only briefly. For all variables a t-test is conducted. The null hypothesis of the means being equal should be rejected since this would indicate a statistically significant difference in the means of participation and non-participation. The smaller the p-value, the stronger the evidence against the null hypothesis and hence all null hypotheses, with exception of age, are rejected. Age is the only variable which cannot be rejected, however, in the multivariate analysis age is not separated into groups but measured in years.

For gender both the main research sample and the subgroup provide evidence that the chance of participating in the stock market is higher for males than for females. Health and stock market participation are negatively related since a decrease in health score indicates a better health condition. Wealth is a relevant variable in explaining the likelihood of stock market participation as well and holding above country average levels of wealth increases the chance of total stock market participation with 18.4% compared to holding below country average levels of wealth. Additionally, education increases the likelihood of stock market participation and so does someone's willingness to take on risk. In other words, having a strong aversion towards risk is accompanied by a decrease in stock market participation.

Figure 3
Social connectedness per European country. This figure provides an overview of average social connectedness per country with 0 and 20 being minimum and maximum respectively. Numbers are based on Wave 6 and Wave 5 of the SHARE database.



As mentioned earlier, for the subgroup direct stock market participation, individuals have to comply with the sole requirement of participating in the stock market via direct means. An individual from the total stock market participation sample is considered participant if he or she participates via direct means, via indirect means, or via both means. Since numerous participants in the subgroup also participate in the stock market via indirect means, the majority of both groups consist of the same participants and the actual difference is in the criterion for being recognized as stock market participant. Total stock market participation rates and direct stock market participation rates are visualized in a single figure provided in Appendix A.3.

5.2 Multivariate results on stock market participation

As mentioned in section 3.3, a coefficient (β) produced by a binary logistic regression cannot be interpreted as a regular slope which directly explains a 1-unit increase in the independent variable, in this research stock market participation. Since the probability of participating in the stock market is not constant at any marginal change, the odds-ratio (Exp (β)) provides the change in odds of stock market participation at any marginal change. Therefore, β is used for statistical significance and implications and Exp (β) is used for economic significance and implications. For a complete overview of all variables used in the remainder of this section, see Appendix A.1.

5.2.1 Financial literacy and stock market participation

To explain a moderating effect on a proposed relation, the relation itself should be tested in advance. Estimates regarding the logistic regression are provided in Table 6. Table 6 is divided into two parts: the left part contains the total stock market participation sample (Total SMP) and the right part contains the direct participation subgroup (Direct SMP). As explained before, the first hypothesis aims at confirming the positive relation between financial literacy (FLT) and stock market participation (SMP). For Total SMP, FLT has a β of 0.352 and this coefficient is statistically significant at the 1% level. Exp (β) is 1.421 and this odds-ratio describes the increase in odds of stock market participation by a 1-unit increase in financial literacy. Hence, for an individual having a financial literacy score of 2, the odds of stock market participation are 1.421 times greater compared to an individual having a financial literacy score of 1. The lowest and highest score for financial literacy are 1 and 5 respectively and therefore 4 intervals in financial literacy is the maximum captured by this model. This indicates, for an individual having a financial literacy score of 5, the odds of stock market participation are 5.684 times greater compared to an individual having a financial literacy score of 1. For Direct SMP, FLT has a β of 0.314 and this coefficient is statistically significant at the 1% level. Exp (β) is 1.369 and this indicates that for an individual having a financial literacy score of 5, the odds of stock market participation are 5.476 times greater compared to an individual having a financial literacy score of 1. Hence, for the subgroup Direct SMP the economic significance of FLT on SMP is slightly less pronounced than for Total SMP. Finally, the control variables do not substantially differ in a statistical or economic sense between Total SMP and Direct SMP. Moreover, all control variables correspond to their proposed relation towards SMP.

Table 6

Multivariate statistics. This table provides a multivariate analysis on stock market participation based on binary logistic regression coefficients. Total SMP consists of 16,918 observations, whereof 2,649 participate and 14,269 do not participate in the stock market, participation rate is 0.157. Direct SMP consists of 16,901 observations, whereof 1,827 participate and 15,074 do not participate in the stock market, participation rate is 0.108. Standard errors provided in parentheses. ***, **, and * indicate the 1%, 5%, and 10% significance level respectively. β is regression coefficient, Exp (β) represents the marginal change. Numbers are based on Wave 6 and Wave 5 of the SHARE database. For complete variable overview see Appendix A.1.

	Total SMP				**Direct SMP**			
	Hypothesis 1		*Hypothesis 2*		*Hypothesis 1*		*Hypothesis 2*	
	β	Exp (β)	β	Exp (β)	β	Exp (β)	β	Exp (β)
Constant	-2.643	0.071	-18.887	0.000	-2.928	0.053	-18.159	0.000
	(0.236)***		(1.369)***		(0.268)***		(1.603)***	
FLT	0.352	1.421	2.538	12.649	0.314	1.369	2.332	10.295
	(0.026)***		(0.340)***		(0.031)***		(0.397)***	
SCN			1.647	5.190			1.544	4.684
			(0.132)***				(0.154)***	
FLT * SCN			-0.221	0.802			-0.204	0.816
			(0.033)***				(0.038)***	
AGE	0.035	1.036	0.030	1.031	0.034	1.034	0.029	1.029
	(0.002)***		(0.003)***		(0.003)***		(0.003)***	
GEN	0.199	1.221	0.217	1.243	0.150	1.162	0.158	1.172
	(0.048)***		(0.049)***		(0.056)***		(0.057)***	
HEA	-0.294	0.745	-0.263	0.768	-0.297	0.743	-0.261	0.770
	(0.023)***		(0.024)***		(0.027)***		(0.027)***	
WEA	0.256	1.292	0.300	1.350	0.248	1.282	0.275	1.317
	(0.016)***		(0.017)***		(0.016)***		(0.017)***	
EDU	0.066	1.068	0.050	1.051	0.061	1.063	0.047	1.048
	(0.006)***		(0.006)***		(0.007)***		(0.007)***	
RIA	-0.838	0.433	-0.787	0.455	-0.808	0.446	-0.758	0.468
	(0.032)***		(0.034)***		(0.035)***		(0.036)***	
-2 log likelihood	12110.323		11307.768		9688.950		9155.862	
Cox and Snell's R^2	0.141		0.181		0.106		0.134	
Number of observations	16,918		16,918		16,901		16,901	

Summarized, for both the Total SMP sample and the Direct SMP subgroup financial literacy has a positive and statistically significant effect on stock market participation. Therefore, the first hypothesis is confirmed. Testing the moderating effect of country-specific social connectedness on the relation between financial literacy and stock market participation will be on reasonable grounds.

5.2.2 Country-specific social connectedness and stock market participation

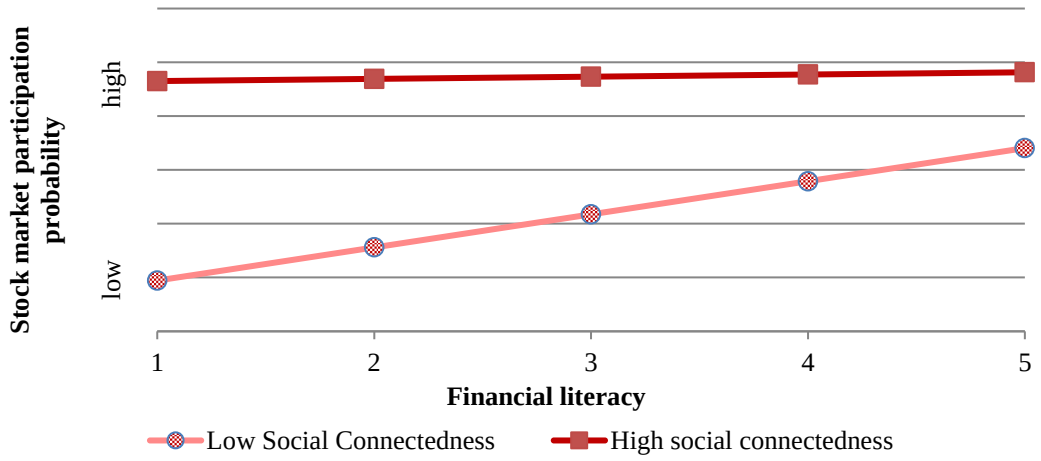
Similar to the first hypothesis, for testing the second hypothesis the regression results provided in Table 6 will be used. As explained before, the second hypothesis aims at confirming a moderating effect of country-specific social connectedness (SCN) on the positive relation between financial literacy (FLT) and stock market participation (SMP). For Total SMP, FLT has a β of 2.538 and this coefficient is statistically significant at the 1% level. SCN has a β of 1.647 and this coefficient is statistically significant at the 1% level. FLT*SCN has a β of -0.221 and the coefficient of the interaction term is significant at the 1% level as well. A comparison between the first and second hypothesis shows that the coefficient FLT remains positive towards SMP. Furthermore, the coefficient SCN is positive and the coefficient of FLT*SCN is negative. For Direct SMP, FLT has a β of 2.332 and this coefficient is statistically significant at the 1% level. The coefficient SCN has a β of 1.544 and is statistically significant at the 1% level. FLT*SCN has a β of -0.204 and this coefficient of the interaction term is significant at the 1% level as well. The percentage stock market participants is 15.7% for the Total SMP sample compared to 10.8% for the Direct SMP subgroup. Albeit less pronounced for Direct SMP, the overall pattern is similar to Total SMP.

A statistical explanation is straightforward since all coefficient are highly significant and thus inferences could be made based on the statistical significance. However, the economic explanation and interpretation of the coefficients is much harder. To fully understand the effect of the interaction term FLT*SCN on the dependent variable SMP, a textual example is provided supplemented with a visualization in Fig. 4. This example takes two random individuals. One individual resides in a high social connectedness country and one individual resides in a low social connectedness country, respectively Italy (8.72) and Austria (11.31) derived from Table 2. If both lines, presented in Fig. 4, would represent one individual, the upper line would be an individual from Austria and the lower line would be an individual of Italy. Assumed is that both individuals are completely equal to each other with respect to AGE, GEN, HEA, WEA, EDU, and RIA. The equality of these variables enables a direct comparison only taking into account

the effect of FLT, SCN, and FLT*SCN on SMP. On the horizontal axis is FLT and on the vertical axis is the chance of SMP. As said before, FLT has a β of 2.538, SCN has a β of 1.647, and FLT*SCN has a β of -0.221. As can be seen in Fig. 4 and supported by the regression coefficients derived from Table 6, apart from each other both FLT and SCN are positively related to SMP. The interaction term is negative and this indicates mathematically speaking, if SCN = 8.72 and FLT = 1, a 1-unit increase in FLT results in 2.538 + (-0.221*8.72) = 0.61. If SCN = 11.31 and FLT = 1, a 1-unit increase in FLT results in 2.538 + (-0.221*11.31) = 0.04. The outcomes are relative, however for the low social connectedness country, a 1-unit increase causes a larger increase in the stock market participation probability than a 1-unit increase in FLT for a high social connectedness country will bring about. In other words, the effect of FLT on SMP decreases when SCN increases. Referring to the example, if both individuals have a 1-unit increase in FLT, the chance of stock market participation will be substantially more for the Italian individual than for an Austrian individual.

Figure 4

Moderating effect of social connectedness. This figure visualizes plotted regression coefficients showing the negative moderating effect of country-specific social connectedness on the probability of stock market participation. Low social connectedness with the lowest country average observed being 8.7 and High social connectedness with the highest country average observed being 11.3. Numbers are based on Wave 6 and Wave 5 of the SHARE database.



The control variables in both Total SMP and Direct SMP have the same direction as expected and are roughly similar to the control variables under the first hypothesis. A minor but significant positive effect is observed for age. Older people have a slightly higher chance in participating in the stock market compared to younger people. Furthermore, being a male increases the chance of stock market participation. Since an increase in health score corresponds to a poorer health, the negative coefficient indicates that a better health increases the chance of participating in the stock market. Wealth is positively related to stock market participation as

is education. Lastly, the relation between risk aversion and stock market participation is negative and thereby confirming that an individual being more risk averse is less likely to participate in the stock market.

In determining whether adding the interaction term to the established relation tested in hypothesis 1 increases the goodness of fit of the ultimate model, the model of hypothesis 2 should be compared to the model of hypothesis 1. The -2 log likelihood presents the deviance or the 'badness of fit' of the model. Interpreting the number on its own is hard, however when the deviance becomes smaller after adding variables the model improves as a consequence of adding extra variables. The observed decrease in the -2 log likelihood of hypothesis 2 compared to hypothesis 1 indicates that the model improves when adding SCN and FLT*SCN. Hence, an increase in Cox and Snell's R^2 is observed since Cox and Snell's R^2 is a pseudo R^2 based on the log likelihood and used for binary logistic regressions. These findings apply to both Total SMP and Direct SMP. Since Cox and Snell's R^2 is a relative measure, this measure is most suitable when comparing the Total SMP sample to the Direct SMP subgroup. For both the first and the second hypothesis the Total SMP sample, which consists of a more comprehensive measure of stock market participation, is best in explaining the variance of the dependent variable SMP.

Summarized, referring to both statistical and economic significance, country-specific social connectedness moderates the relation between financial literacy and stock market participation. This indicates that financial literacy becomes less (more) influential in explaining the stock market participation in countries with higher (lower) social connectedness. Hence, the second hypothesis is confirmed.

5.3 Robustness tests

The dependent variable in this research is not measured by a proxy which could be simply substituted for another proxy and therefore robustness testing with another dependent variable would not be so targeted. Nonetheless, the subgroup direct stock market participation focusses on one specific requirements and therefore also could serve as a check to total stock market participation. With reference to the independent variables, the main contribution of this research is the country-specific social connectedness and hence a robustness check using another database containing macroeconomic data would increase the reliability of the results and the confirmation of the second hypothesis. Klein (2013) suggests that social cohesion is

developed and maintained at the society's level and hence social cohesion is a rather macroeconomic factor. The Organization for Economic Co-operation and Development (OECD) presents social cohesion indicators referring to charity donation, social support to strangers, and volunteering in an organization (OECD, 2014). The variable country-specific social cohesion (SCH) uses combined data on the aforementioned social cohesion indicators. The indicators are equally weighted and are considered nationally representative (OECD, 2014). Social cohesion is provided as a fraction of a country's population and this measure is available for a large number of countries all over the world. A figure showing the fraction of social cohesion per country is provided in Appendix A.4. With regards to this research, data for Croatia is missing and therefore Croatia is excluded in robustness testing. The robustness tests are presented in Table 7.

Table 7

Robustness tests. This table provides a multivariate analysis serving as robustness check on stock market participation based on binary logistic regression coefficients. Total SMP consists of 15,376 observations, whereof 2,585 participate and 12,791 do not participate in the stock market, participation rate is 0.168. Direct SMP consists of 15,360 observations, whereof 1,774 participate and 13,586 do not participate in the stock market, participation rate is 0.115. Standard errors provided in parentheses. ***, **, and * indicate the 1%, 5%, and 10% significance level respectively. β is regression coefficient, Exp (β) represents the marginal change. Numbers are based on the OECD database and on Wave 6 and Wave 5 of the SHARE database. For complete variable overview see Appendix A.1.

	Total SMP		**Direct SMP**	
	β	Exp (β)	β	Exp (β)
Constant	-7.196	0.001	-7.082	0.001
	(0.729)***		(0.839)***	
FLT	0.780	2.181	0.680	1.989
	(0.173)***		(0.200)***	
SCH	0.137	1.147	0.125	1.133
	(0.019)***		(0.021)***	
FLT * SCH	-0.013	0.987	-0.011	0.989
	(0.005)***		(0.005)**	
AGE	0.029	1.029	0.028	1.029
	(0.003)***		(0.003)***	
GEN	0.211	1.235	0.147	1.159
	(0.050)***		(0.058)**	
HEA	-0.252	0.777	-0.251	0.778
	(0.024)***		(0.028)***	
WEA	0.337	1.401	0.308	1.361
	(0.019)***		(0.019)***	
EDU	0.050	1.052	0.047	1.049
	(0.006)***		(0.007)***	
RIA	-0.792	0.453	-0.763	0.466
	(0.034)***		(0.037)***	
-2 log likelihood	10919.954		8842.428	
Cox and Snell's R^2	0.178		0.131	
Number of observations	15,376		15,360	

The regression results in Table 7 show that all coefficients keep the same direction and remain significant compared to the regression results in Table 6. Both FLT and SCH remain positive whereas FLT*SCH remains negative indicating that the moderating effect of SCN on SMP is robust. For the Direct SMP subgroup, FLT*SCH becomes slightly less significant although the negative moderating effect remains. Since the moderating effect of country-specific social connectedness on the relation between financial literacy and stock market participation is the main contribution of this research, the results of this research are considered robust.

Additionally, the results for Total SMP are considered robust if the independent variable FLT is replaced by another independent variable serial (SER). The variable SER measures an individual's ability to subtract a number over and over from the previous number and therefore SER is another proxy for mathematical and cognitive abilities. A figure showing the serial score per country is provided in Appendix A.5. Although SER is not financially framed, Lusardi et al. (2010) suggest that the cognitive ability of an individual is a powerful predictor of an individual's financial literacy. Appendix A.6. presents the robustness tests for FLT and shows that for the Total SMP sample, the results remain intact. Appendix A.7. presents the robustness tests if all main variables in this research are replaced at the same time, thus FLT is replaced by SER, SCN is replaced by SCH, and FLT*SCN is replaced by SER*SCH. Only the coefficients in the Total SMP sample remain significant, albeit to a lesser extent both in a statistical and economic sense.

6. Conclusion and discussion

6.1 Conclusions and implications

Until today, the stock market participation puzzle is relatively incomplete. Despite its proven importance, scholars still could not completely rationalize the rather low stock market participation rates and the differences in stock market participation rates between countries. Traditional research has mainly focused on individual characteristics with regards to stock market participation. The first hypothesis of this research builds upon a large and extended base of existing literature and, while taking into account more countries, this research reconfirms that financial literacy positively influences stock market participation. The results for both

samples indicate that the likelihood of stock market participation increases by an increase in financial literacy. This conclusion is in line with Van Rooij et al. (2011, 2012).

Contemporary literature refers to the presence of broader, institutional factors explaining the widely dispersed stock market participation rates. Cross-country differences are brought forward by authors as Christelis et al. (2010), and Jappelli (2010). The second hypothesis draws from Thomas and Spataro (2015), whom indicate that country-level institutional factors ultimately act as determinants for stock market participation. The main contribution of this research is to provide insight into country-specific social connectedness moderating the relation between financial literacy and stock market participation. Although this research reveals a significant, positive, and direct relation between country-specific social connectedness and stock market participation, not particularly the direct effect but the moderating effect of country-specific social connectedness on stock market participation is of great importance. As indicated in the academic literature noted in Section 2.2, individuals shape their portfolios and base their financial decisions depending on what they are capable of (Christelis et al., 2010). Furthermore, stock market participation is dependent on the cost of entry (Alan, 2006) and for individuals being financially illiterate, the stock market entry and participation costs are higher. Hence, financial literacy should be less relevant in countries where information required to participate in the stock market can easily be obtained from an individual's surroundings. The results show a negative relation between the moderating effect of country-specific social connectedness and stock market participation. The relation being negative statistically indicates that the effect of financial literacy on stock market participation decreases (increases) when a country's social connectedness increases (decreases). The relation being negative in an economic sense implies that a change in financial literacy has less impact on stock market participation in a high social connectedness country than in a low social connectedness country. This suggests that social connectedness partly substitutes financial literacy in the stock market participation decision. Consequently, and referring to the second hypothesis, a country's level of social connectedness moderates the relation between financial literacy and stock market participation.

In this research, the total stock market participation sample is most comprehensive and consists of both direct and/or indirect stock market participants. Besides, the subgroup direct participation consists of individuals with the sole requirement of direct stock market participation. Both samples succeed in confirming that financial literacy positively influences stock market participation and county-specific social connectedness moderates the

aforementioned relation. For both samples the relations remain intact if country-specific social connectedness is replaced by country-specific social cohesion. Additional robustness tests indicate that only the results of the total participation sample remain intact, however since the total participation sample is most comprehensive in its definition of a stock market participant, the findings of this research are considered robust.

Referring to the research question, the predictive power of financial literacy on stock market participation decreases if country-specific social connectedness increases. Contrary, financial literacy becomes a more important determinant for stock market participation when a country's social connectedness decreases. These findings implicate that it is particularly interesting for policy makers of low social connectedness countries to invest in increasing their country's financial literacy. This investment will yield a relatively large increase in stock market participation compared to a similar financial literacy increase in a very socially connected country. Moreover, an increase in stock market participation will eventually benefit society (Cocco et al., 2005).

6.2 Limitations and future research

This research offers new significant relations by using recent multi-country data and the newly introduced computed variable social connectedness derived from the SHARE database. Unfortunately, the database mainly contains respondents aged 50 years and older. Despite the older population controlling a substantial part of the financial resources, it remains a skewed reflection of reality. Furthermore, in this research, stock market participation in its most comprehensive form, the total participation sample, consists of both direct and/or indirect stock market participants. Considerable caution is required in defining an indirect stock market participant since the effects of financial literacy could fade away under certain circumstances. An example could relate to a wealthy individual requesting an external agent to manage its capital. The individual can be completely financially illiterate and still argue that he or she participates in the stock market, albeit indirectly via an external agent. The measure indirect stock market participation can thus possibly bypass the effect of financial literacy and harm the results of the study conclusively.

Although the creation of the subgroup direct stock market participants is possible, the creation of a subgroup indirect stock market participants is not possible with the current data because the subgroup is not sufficiently large to draw conclusions from. To sharpen the contrast between investors of both groups, future research could attempt to make a group of participants

only participating via direct means, a group only participating via indirect means, and a group participating via both means. However, as noted before, this realization depends on data availability. Fortunately, SHARE announced the release of Wave 7 which will become available in spring 2019. Wave 7 entails a full coverage of the 28 member-countries in the European Union. When research is conducted with data derived from Wave 7, comparisons across all EU members-countries are possible which could result in even more detailed country-specific inferences. Results from such research could be even more interesting to policy makers, politicians, managers in monetary economic positions, and persons in managerial positions within banks. All in all, deepening the understanding of the determinants of stock market participation could result in taking appropriate action which eventually benefits society and a country's welfare.

7. References

Alan, S., 2006. Entry costs and stock market participation over the life cycle. Review of Economic Dynamics 9, 588-611.

Almenberg, J., Dreber, A., 2015. Gender, stock market participation and financial literacy. Economics Letters 137, 140-142.

Bailey, M., Cao, R., Kuchler, T., Stroebel, J., Wong, A., 2017. Measuring social connectedness. Unpublished working paper. National Bureau of Economic Research.

Baker, H., Ricciardi, V., 2014. Investor Behavior: The Psychology of Financial Planning and Investing. John Wiley and Sons, New Jersey.

Balloch, A., Nicolae, A., Philip, D., 2014. Stock market literacy, trust, and participation. Review of Finance 19, 1925-1963.

Börsch-Supan, A., Brandt, M., Hunkler, C., Kneip, T., Korbmacher, J., Malter, F., Schaan, B., Stuck, S., Zuber, S., 2013. Data resource profile: the Survey of Health, Ageing and Retirement in Europe (SHARE). International Journal of Epidemiology 42, 992-1001.

Börsch-Supan, A., 2017. Survey of Health, Ageing and Retirement in Europe (SHARE) Wave 6. Release version: 6.0.0. SHARE-ERIC. Data set.

Brown, J., Ivković, Z., Smith, P., Weisbenner, S., 2008. Neighbors matter: causal community effects and stock market participation. Journal of Finance 63, 1509-1531.

Campbell, J., 2006. Household finance. Journal of Finance 61, 1553-1604.

Changwony, F., Campbell, K., Tabner, I., 2015. Social engagement and stock market participation. Review of Finance 19, 317-366.

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

Cocco, J., Gomes, F., Maenhout, P., 2005. Consumption and portfolio choice over the life cycle. Review of Financial Studies 18, 491-533.

Courvoisier, D., Combescure, C., Agoritsas, T., Gayet-Ageron, A., Perneger, T., 2011. Performance of logistic regression modeling: beyond the number of events per variable, the role of data structure. Journal of Clinical Epidemiology 64, 993-1000.

Dewey, M., Prince M., 2005. Cognitive function. In: Börsch-Supan, A., Brugiavini, A., Jürges, H., Mackenbach, J., Siegriest, J., Weber, G., eds. Health, Aging and Retirement in Europe: First Results from the Survey of Health, Aging and Retirement in Europe, Mannheim.

Duflo, E., Saez, E., 2002. Participation and investment decisions in a retirement plan: the influence of colleagues' choices. Journal of Public Economics 85, 121-148.

Georgarakos, D., Pasini, G., 2011. Trust, sociability, and stock market participation. Review of Finance 15, 693-725.

Graham, J., Harvey, C., Huang, H., 2009. Investor competence, trading frequency, and home bias. Management Science 55, 1094-1106.

Granovetter, M., 2005. The impact of social structure on economic outcomes. Journal of Economic Perspectives 19, 33-50.

Guiso, L., Haliassos, M., Jappelli, T., 2003. Household stockholding in Europe: where do we stand and where do we go? Economic Policy 18, 123-170.

Guiso, L., Jappelli, T., 2005. Awareness and stock market participation. Review of Finance 9, 537-567.

Haliassos, M., Bertaut, C., 1995. Why do so few hold stocks? Economic Journal 105, 1110-1129.

Han, B., Hirshleifer, D., 2016. Social transmission bias and active investing. Unpublished working paper. Rotman School of Management.

Hong, H., Kubik, J., Stein, J., 2004. Social interaction and stock-market participation. Journal of Finance 59, 137-163.

Hong, H., Kubik, J., Stein, J., 2005. Thy neighbor's portfolio: word-of-mouth effects in the holdings and trades of money managers. Journal of Finance 60, 2801-2824.

Jappelli, T., 2010. Economic literacy: an international comparison. Economic Journal 120, 429-451.

Jappelli, T., Padula, M., 2013. Investment in financial literacy and saving decisions. Journal of Banking and Finance 37, 2779-2792.

Kaustia, M., Knüpfer, S., 2012. Peer performance and stock market entry. Journal of Financial Economics 104, 321-338.

Kiss, A., Danis, W., 2008. Country institutional context, social networks, and new venture internationalization speed. European Management Journal 26, 388-399.

Klapper, L., Lusardi, A., Panos, G., 2013. Financial literacy and its consequences: evidence from Russia during the financial crisis. Journal of Banking and Finance 37, 3904-3923.

Klein, C., 2013. Social capital or social cohesion: what matters for subjective well-being? Social Indicators Research 110, 891-911.

Lusardi, A., Mitchell, O., 2007. Baby boomer retirement security: the roles of planning, financial literacy, and housing wealth. Journal of Monetary Economics 54, 205-224.

Lusardi, A., Mitchell, O., Curto, V., 2010. Financial literacy among the young. Journal of Consumer Affairs 44, 358-380.

Mandell, L., Klein, L., 2007. Motivation and financial literacy. Financial Services Review 16, 105-116.

Mankiw, N., Zeldes, S., 1991. The consumption of stockholders and nonstockholders. Journal of Financial Economics 29, 97-112.

McArdle, J., Smith, J., Willis, R., 2009. Cognition and economic outcomes in the health and retirement survey. Unpublished working paper. National Bureau of Economic Research.

Mehra, R., Prescott, E., 1985. The equity premium: a puzzle. Journal of Monetary Economics 15, 145-161.

OECD, 2014. Helping others. In: Society at a Glance 2014: OECD Social Indicators. OECD Publishing, Paris.

Peduzzi, P., Concato, J., Kemper, E., Holford, T., Feinstein, A., 1996. A simulation study of the number of events per variable in logistic regression analysis. Journal of Clinical Epidemiology 49, 1373-1379.

Rosen, H., Wu, S., 2004. Portfolio choice and health status. Journal of Financial Economics 72, 457-484.

Thomas, A., Spataro, L., 2015. Financial literacy, human capital and stock market participation in Europe: an empirical exercise under endogenous framework. Unpublished working paper. Economic Sciences. University of Pisa.

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

Van Rooij, M., Lusardi, A., Alessie, R., 2012. Financial literacy, retirement planning and household wealth. Economic Journal 122, 449-478.

Vissing-Jørgensen, A., Attanasio, O., 2003. Stock-market participation, intertemporal substitution, and risk-aversion. American Economic Review 93, 383-391.

Vittinghoff, E., McCulloch, C., 2007. Relaxing the rule of ten events per variable in logistic and Cox regression. American Journal of Epidemiology 165, 710-718.

8. Appendix

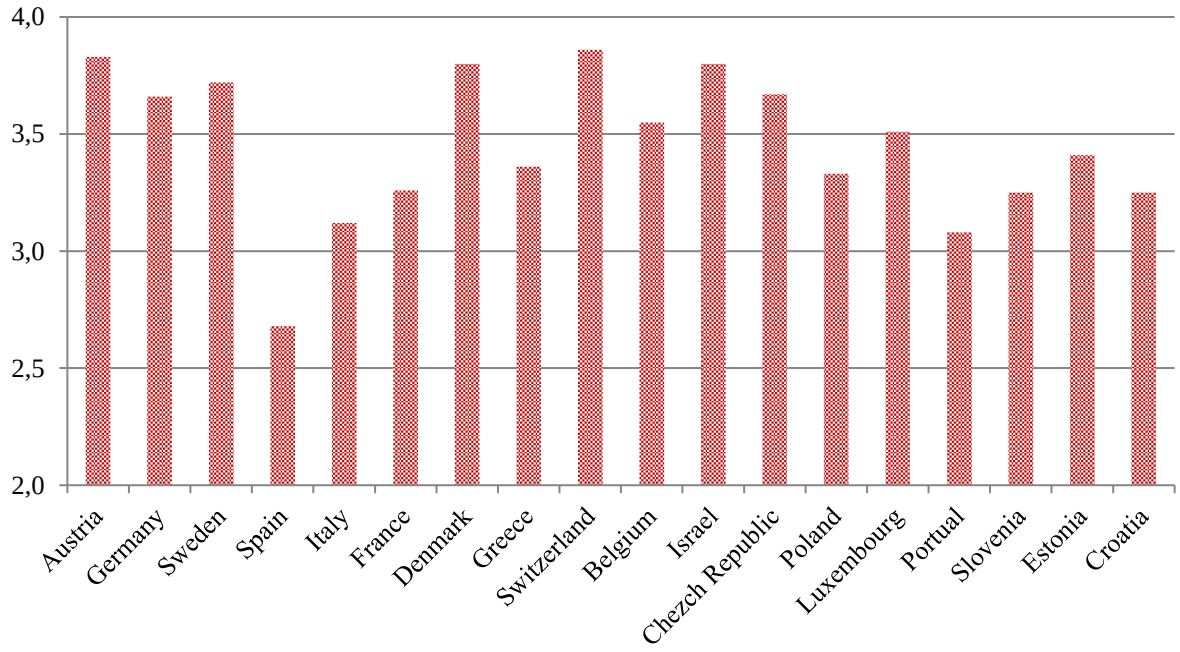
Appendix A.1.

Variable overview. This table provides an overview of the dependent variable, independent variables, control variables, and variables used for the robustness tests. All variables are derived from the SHARE database Wave 6 and Wave 5 with exception of social cohesion (SCH) which is derived from the OECD.

SMP	Stock market participation
FLT	Financial literacy
SCN	Country-specific social connectedness
AGE	Age
GEN	Gender
HEA	Health
WEA	Wealth
EDU	Education
RIA	Risk aversion
SCH	Country-specific social cohesion
SER	Serial

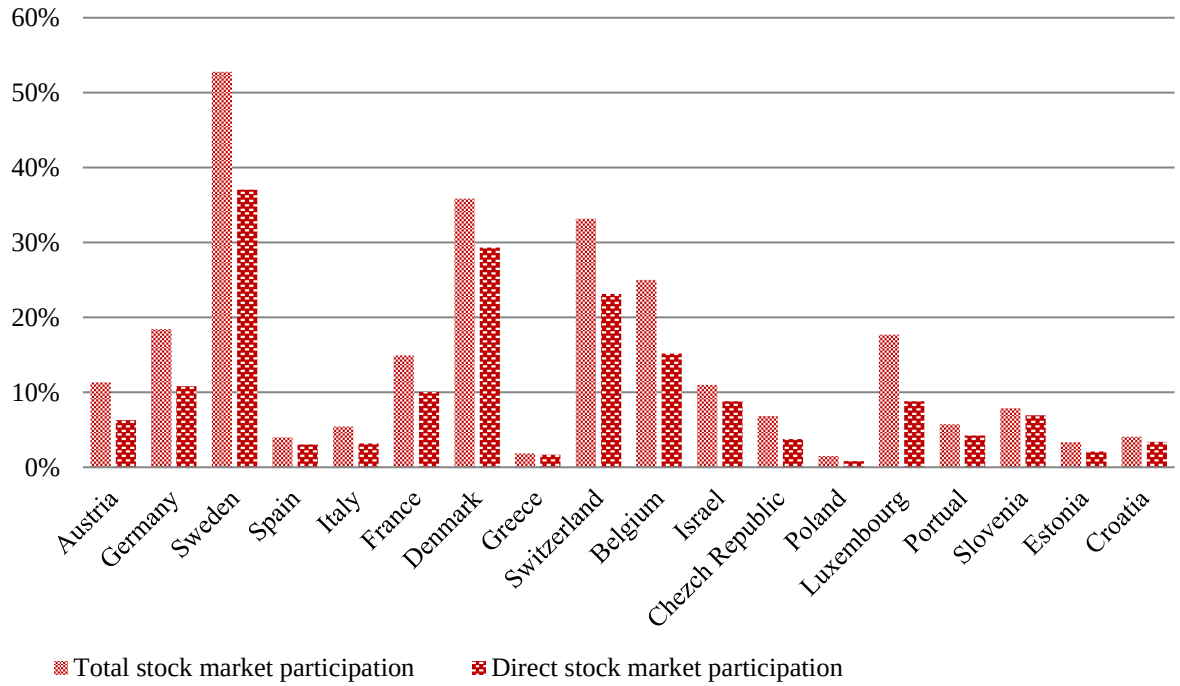
Appendix A.2.

Financial literacy score per European country. This figure provides an overview of the average financial literacy score per country with 1 and 5 being minimum and maximum respectively. Numbers are based on Wave 6 and Wave 5 of the SHARE database. Number of observations is 16,918.



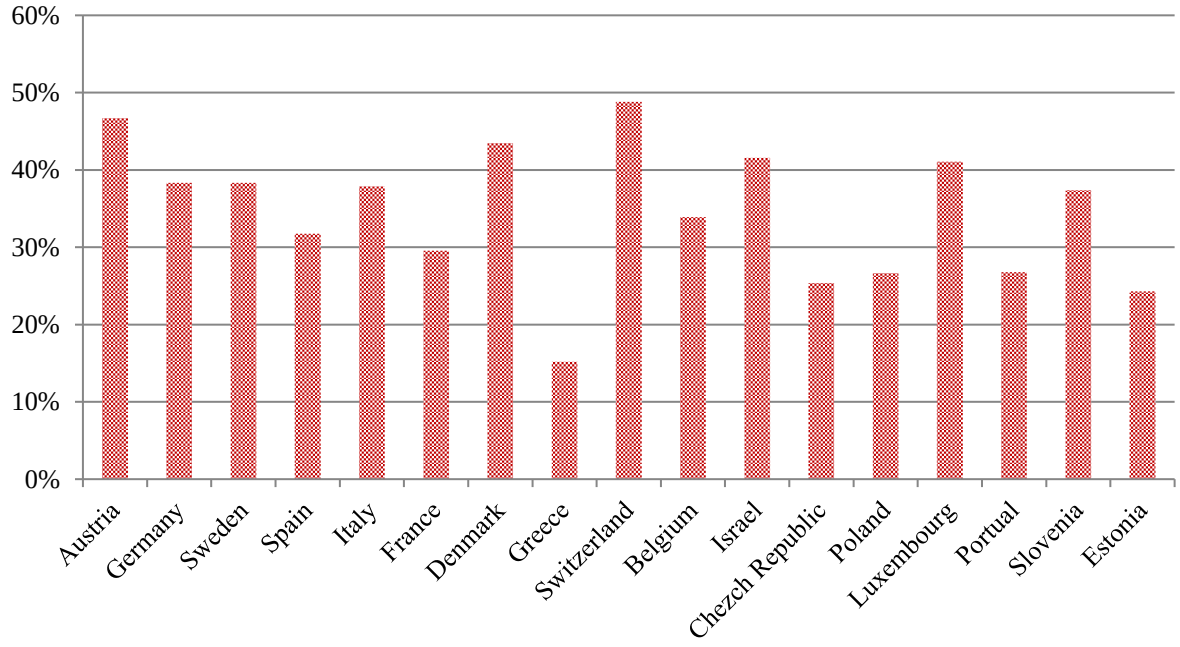
Appendix A.3.

Total stock market participation rates and direct stock market participation rates per European country. This figure provides an overview of the difference between total stock market participation as fraction of the population per country and direct stock market participation as fraction of the population per country. Numbers are derived from Wave 6 of the SHARE database.



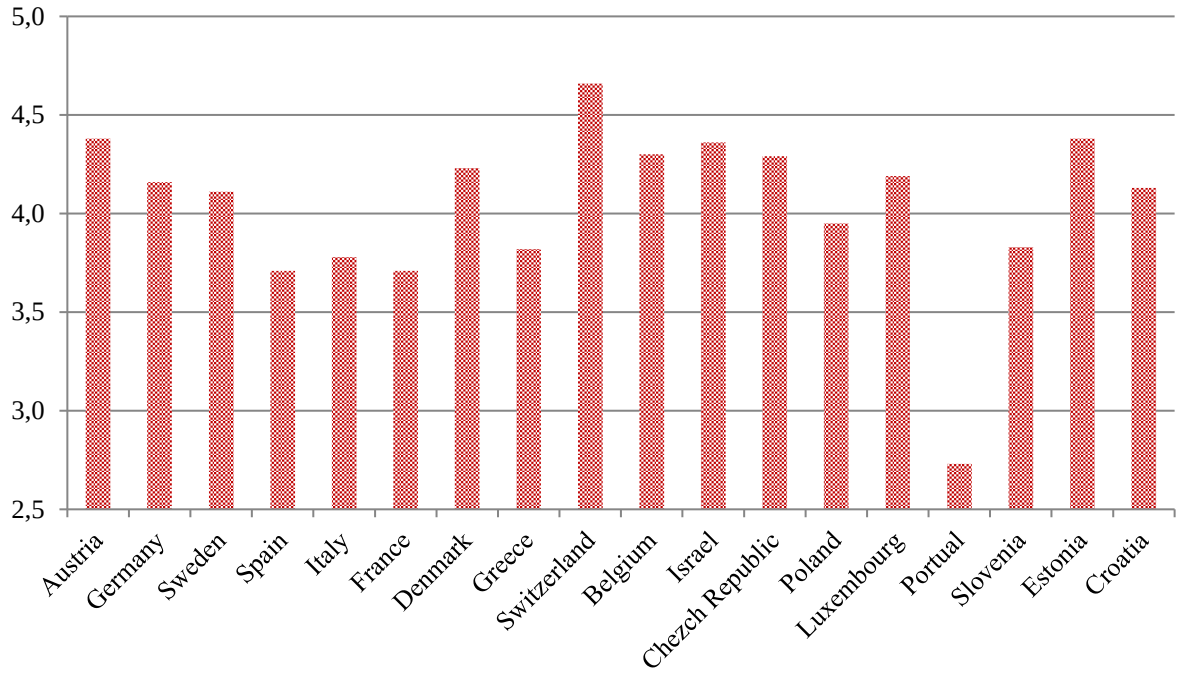
Appendix A.4.

Social cohesion per European country. This figure provides an overview of average social cohesion as fraction of the population per country derived from charity donation, social support to strangers, and volunteering in an organization. Numbers are derived from OECD database.



Appendix A.5.

Serial score per European country. This figure provides an overview of average serial score per country with 1 and 5 being minimum and maximum respectively. Numbers are based on Wave 6 of the SHARE database. Number of observations is 15,679.



Appendix A.6.

Robustness tests. This table provides a multivariate analysis serving as robustness check on stock market participation based on binary logistic regression coefficients. Total SMP consists of 15,679 observations, whereof 2,611 participate and 13,068 do not participate in the stock market, participation rate is 0.167. Direct SMP consists of 15,662 observations, whereof 1,797 participate and 13,865 do not participate in the stock market, participation rate is 0.115. Standard errors provided in parentheses. ***, **, and * indicate the 1%, 5%, and 10% significance level respectively. β is regression coefficient, Exp (β) represents the marginal change. Numbers are based on Wave 6 and Wave 5 of the SHARE database. For complete variable overview see Appendix A.1.

	Total SMP		**Direct SMP**	
	β	Exp (β)	β	Exp (β)
Constant	-10.600	0.000	-9.479	0.000
	(0.687)***		(0.780)***	
SER	1.547	4.695	0.028	1.029
	(0.744)**		(0.857)	
SCN	0.920	2.510	0.774	2.169
	(0.064)***		(0.073)***	
SER * SCN	-0.147	0.863	-0.001	0.999
	(0.073)**		(0.083)	
AGE	0.031	1.031	0.029	1.029
	(0.003)***		(0.003)***	
GEN	0.282	1.325	0.217	1.242
	(0.049)***		(0.057)***	
HEA	-0.275	0.759	-0.273	0.761
	(0.024)***		(0.027)***	
WEA	0.312	1.367	0.284	1.329
	(0.017)***		(0.017)***	
EDU	0.062	1.064	0.058	1.060
	(0.006)***		(0.007)***	
RIA	-0.800	0.449	-0.773	0.462
	(0.033)***		(0.036)***	
-2 log likelihood	11143.669		8992.977	
Cox and Snell's R^2	0.173		0.129	
Number of observations	15,679		15,662	

Appendix A.7.

Robustness tests. This table provides a multivariate analysis serving as robustness check on stock market participation based on binary logistic regression coefficients. Total SMP consists of 14,237 observations, whereof 2,548 participate and 11,689 do not participate in the stock market, participation rate is 0.179. Direct SMP consists of 14,221 observations, whereof 1,745 participate and 12,476 do not participate in the stock market, participation rate is 0.123. Standard errors provided in parentheses. ***, **, and * indicate the 1%, 5%, and 10% significance level respectively. β is regression coefficient, Exp (β) represents the marginal change. Numbers are based on the OECD database and on Wave 6 and Wave 5 of the SHARE database. For complete variable overview see Appendix A.1.

	Total SMP		**Direct SMP**	
	β	Exp (β)	β	Exp (β)
Constant	-5.038	0.006	-4.905	0.007
	(0.425)***		(0.476)***	
SER	0.750	2.116	0.330	1.392
	(0.404)*		(0.457)	
SCH	0.104	1.110	0.089	1.093
	(0.010)***		(0.011)***	
SER * SCH	-0.018	0.982	-0.007	0.993
	(0.011)*		(0.012)	
AGE	0.030	1.030	0.029	1.029
	(0.003)***		(0.003)***	
GEN	0.280	1.323	0.211	1.235
	(0.050)***		(0.058)**	
HEA	-0.260	0.771	-0.259	0.772
	(0.025)***		(0.028)***	
WEA	0.351	1.420	0.318	1.374
	(0.019)***		(0.019)***	
EDU	0.063	1.065	0.058	1.060
	(0.006)***		(0.007)***	
RIA	-0.811	0.444	-0.781	0.458
	(0.034)***		(0.037)***	
-2 log likelihood	10712.420		8658.686	
Cox and Snell's R^2	0.171		0.127	
Number of observations	14,237		14,221	