PhD THESIS DECLARATION
(use a pc to fill it out)

I, the undersigned

FAMILY NAME	Balakina
NAME	Olga
Student ID no.	1612431

Thesis title:

Essays in Social Finance and Banking

PhD in	Economics and Finance
Cycle	27
Student's Tutor	Nicola Gennaioli
Year of thesis defence	2017

DECLARE

under my responsibility:

1) that, according to Italian Republic Presidential Decree no. 445, 28th December 2000, mendacious declarations, falsifying records and the use of false records are punishable under the Italian penal code and related special laws. Should any of the above prove true, all benefits included in this declaration and those of the temporary "embargo" are automatically forfeited from the beginning;

2) that the University has the obligation, according to art. 6, par. 11, Ministerial Decree no. 224, 30th April 1999, to keep a copy of the thesis on deposit at the "Bilioteche Nazionali Centrali" (Italian National Libraries) in Rome and Florence, where consultation will be permitted, unless there is a temporary "embargo" protecting the rights of external bodies and the industrial/commercial exploitation of the thesis;

3) that the Bocconi Library will file the thesis in its “Archivio Istituzionale ad Accesso Aperto” (Institutional Registry) which permits online consultation of the complete text (except in cases of temporary “embargo”);

4) that, in order to file the thesis at the Bocconi Library, the University requires that the thesis be submitted online by the student in unalterable format to Società NORMADEC (acting on behalf of the University), and that NORMADEC will indicate in each footnote the following information:

- PhD thesis (*write thesis title*) Essays in Social Finance and Banking ____;
- by (S*tudent's family name and name)* Balakina Olga ;
- defended at Università Commerciale “Luigi Bocconi” – Milano in the year (year of defence) 2017 ;
- the thesis is protected by the regulations governing copyright (Italian law no. 633, 22nd April 1941 and subsequent modifications). The exception is the right of Università Commerciale “Luigi Bocconi” to reproduce the same, quoting the source, for research and teaching purposes;

- **only when a separate “Embargo” Request has been undersigned**: the thesis is subject to “embargo” for (indicate duration of the “embargo”) .... months;

5) that the copy of the thesis submitted online to Normadec is identical to the copies handed in/sent to the members of the Thesis Board and to any other paper or digital copy deposited at the University offices, and, as a consequence, the University is absolved from any responsibility regarding errors, inaccuracy or omissions in the contents of the thesis;

6) that the contents and organization of the thesis is an original work carried out by the undersigned and does not in any way compromise the rights of third parties (Italian law, no. 633, 22nd April 1941 and subsequent integrations and modifications), including those regarding security of personal details; therefore the University is in any case absolved from any responsibility whatsoever, civil, administrative or penal, and shall be exempt from any requests or claims from third parties;

7) **choose 7a or 7b**:

7a) that the thesis is not subject to “embargo”, i.e. that it is not the result of work included in the regulations governing industrial property; it was not written as part of a project financed by public or private bodies with restrictions on the diffusion of the results; is not subject to patent or protection registrations.

or

7b) that the PhD thesis is subject to “embargo” as per the separate undersigned “PhD Thesis Temporary ″Embargo″ Request”.

Date _____30.01.2017______

ABSTRACT

The following thesis is covering two topics: Social Finance and Banking. The first part of the thesis is study investigating cross-border capital flows in order to verify the existence and direction of the effect of the soft regulation promoted by international organizations against banking secrecy which characterized the so called tax and financial heavens. The study finds evidence that in general the stigma effect doesn't exist. We test whether being included and later excluded from the FATF blacklist is an effective measure that influences countries' cross-border capital flows. The second part of the thesis contains a study which introduces an equilibrium model of stock market participation with a social network to study how information diffusion affect the decision to enter the stock market. We find that connectivity in the social network has a positive impact on stock market participation. Finally, we find that the model with social networks outperform standard models in predicting the stock market participation. The third part of the thesis is a study, in which we construct a model of investment choice in the presence of peer effect and empirically test whether the portfolio choices of individuals are influenced by the choices of their peers using data for the full Danish population from 2000 to 2013. We find that each type of networks has significant impact on the stock market participation when taken separately. In addition, we find that if we consider all the networks in the same model, only colleagues investment behavior significantly affects stock market participation. Furthermore, we see that the relative influence of each network depends on the characteristics of the network and the characteristics of the household.

Bank Secrecy in Offshore Centres and Capital Flows: Does Blacklisting Matter?

Olga Balakina, Angelo D'Andrea and Donato Masciandaro[1]

This study analyses cross-border capital flows in order to verify the existence and direction of the effect of the soft regulation promoted by international organizations against banking secrecy which characterized the so called tax and financial heavens. This effect is called in the literature Stigma Effect. But both the existence and the direction of the stigma effect are far from being obvious: the international capital flows can simply neglect the relevance of the blacklisting, or worst the attractiveness of banking secrecy can produce a race to the bottom: the desire to elude more transparent regulation can sensibly influence the capital movements. We test whether being included and later excluded from the FATF blacklist is an effective measure that influences countries' cross-border capital flows. Using annual panel data for the period 1996-2014, we applied our framework to 126 countries worldwide. We find evidence that in general the stigma effect doesn't exist.

JEL Classification Numbers: F21, K42

Keywords: bank secrecy, offshore centres, international capital flows, name and shame regulation, money laundering

---

[1] Corresponding author: donato.masciandaro@unibocconi.it. The research has been supported by the Baffi Carefin Centre, Bocconi University. The authors warmly thank the anonymous referees for their constructive comments and useful insights on earlier drafts. Donato Masciandaro pleasantly acknowledges that acting as Consultant for the Inter-American Development Bank has been a unique and precious experience to enrich his knowledge on how the blacklisting process concretely interacts with the financial dynamics. The usual disclaimers are applied.

1. <u>Introduction</u>

On April 2016 revelations of the Panama Papers spotlighted the role that banking secrecy - which is offered in the so called tax and financial centres and territories - perform in the global economy. The facts have caused increasing concern that banking secrecy lies at the centre of an international web of illegal and criminal conduct. In parallel, several policymakers in advanced countries have emphasised the need for enforcing the blacklisting tool against the territories that breach transparency standards. But does the blacklisting work?

Banking secrecy is an evergreen issue for the national and international debate. In the aftermath of the Global Financial Crisis, the fight against bank secrecy as well as against tax and financial havens has become a political priority in advanced countries.

It is often the case that international organisations and national governments do not have strong legal instruments to impose strict measures to prevent and combat banking secrecy. For this reason, soft law practices, such as blacklisting, have been introduced. The aim of the soft law tools is to put the investigated country under intense international financial pressures, using the “name and shame” approach. Under the “name and shame” approach, institutional regulatory organizations and/or national governments disclose names of non-compliant countries and/or non-compliant banks to the public, supplementing the disclosure with forms of official opprobrium (Brummer, 2012). This approach is increasingly applied in the international context and it aims to address policy coordination problems among national policymakers and regulators (Greene and Boehm, 2012).

This paper looks at cross-border capital flows the period 1996-2014 in order to verify the existence and the direction of the so-called stigma effect, i.e. the effect of the blacklisting in addressing banking secrecy.

Country compliance with the international standards of the blacklisting policy named Anti-Money Laundering and Combating the Financing of Terrorism – **AML/CFT** thereafter – gained momentum in the national policymaking all around the world in the last two decades.

Established by the Financial Action Task Force (**FATF**) in 1999, nowadays the international standard consists of 49 Recommendations, dealing respectively with anti-money laundering (forty recommendations) and combating terrorist financing (nine recommendations). Since 2000, FATF has periodically issued lists – **Blacklists** thereafter – of Non-Cooperative Countries and Territories (**NCCTs**), which identify the jurisdictions that FATF believes to be non-compliant with international best practices.

In order to prevent and combat illegal financial flows, international organizations do not have hard legal commitments at their disposal; therefore, they resort to blacklisting by FATF as a soft law practice. The aim of the listing procedure is to put Black-Listed Countries (**BLCs**) under intense international financial pressure, by employing the "name and shame" approach in order to produce the so-called ***stigma effect*** (Masciandaro, 2005 and 2008). The stigma effect represents an inverse relationship between blacklisting and international capital flows. Indeed, the event of being blacklisted decreases the international capital flows towards a country. Two sources of pressures on the BLC are expected to work.

On the one side, most countries that interact with a BLC evaluate its financial transactions as suspicious. This occurrence leads to more stringent and costly monitoring procedures. Banks operating in multiple jurisdictions are the most concerned by these **monetary costs**, including compliance costs. The AML/CFT cost of compliance seems to continuously increase, at an average rate of about 45 per cent (KPMG 2011).

Along with monitoring costs, financial transactions with a BLC can imply **reputational costs**. Suspicious financial transactions attract more and more attention from supranational organizations, national policymakers and regulators, and international media. For banking institutions, engagement in opaque financial transactions can increase reputational risks. Just to cite some more recent and meaningful episodes, it is worth mentioning that in 2012-15 various international banks have been investigated for alleged illicit financial transactions and fined, or solicited, to improve their compliance (Powell, 2013). Transactions with BLCs can produce such a kind of negative reputational effects.

Because of the potential damage caused by the stigma effect, international banks may have a strong incentive to avoid business with BLCs.

In the same way, the stigma effect can be considered because of the "name and shame" approach.

However, both the existence and the direction of the stigma effect are far from being obvious. As it was pointed out in previous studies – Masciandaro (2005 and 2008) and Masciandaro et al.

(2007) – the AML/CFT non-compliance of a country can be attractive under specific conditions, such as the potential existence of a worldwide demand for non-transparent financial transactions. A BLC can be attractive for banking and non-banking institutions seeking to promote lightly regulated products and services to their wealthy and/or sophisticated clients. The international banking industry as a whole can have incentives to take advantages from the existence of BLCs.

Therefore, the stigma effect, meant to be "a stick" for all the countries not in compliance with the regulation, can turn out into "a carrot".

The ***stigma paradox*** can emerge. A specific case of regulatory arbitrage that creates the so-called "race to the bottom" strategy, which implies the desire to elude more prudent regulation (Barth et al., 2006) and that can sensibly influence the international capital movements (Houston, 2011).

Finally, a third possibility has to be considered: the behaviour of the international banking institutions in the cross-border business can be simply driven by factors other than the stigma effect (Kurdle 2009). In this case, the ***stigma neutrality*** holds.

The relevance of the stigma effect has become increasingly important in recent times, when policymakers, regulators and scholars seek to understand which institutional, regulatory and historical features can attract or discourage international capital flows (Papaioannu, 2009; Reinhardt et al., 2010; Houston et al., 2011; Qureshi et al., 2011; Milesi Feretti and Tille, 2011; Chitu et al., 2013). The financial effects of regulation are particularly relevant when the AML/CFT rules are under discussion.

This paper aims to empirically evaluate the trend, magnitude and robustness of the stigma effect, by focusing on the impact of the FATF blacklisting on the relationships between international financial flows and the BLCs banking systems.

So far empirical evidence is sparse and mixed (Kurdle 2009, Masciandaro) – with cases of stigma effect, stigma paradox and stigma neutrality being detected – and therefore inconclusive.

To understand the kind of influence that FATF blacklisting can have on BLCs, our research focuses on how international capital flows respond to the stigma signals provided by the FATF. The stigma effect is based on the assumption that blacklisting procedures alter the attractiveness of a country for capital flows. The non-compliance of a country with AML/CFT standards (listing) can decrease the overall amount of financial transactions (volume effect) and/or decrease its efficiency to manage those capitals (cost effect). Of course, the opposite can happen when a country is delisted.

In this paper, we aim to find empirical evidence of whether and to what extent FATF blacklisting affects the volume of financial transactions. The stigma effect, as a signal that enables to distinguish between compliant and non-compliant countries, can have deterring effects, since transactions with non-compliant countries imply higher monitoring and/or reputational costs. Observing the signal, international banks allocate their activities accordingly. The effects of blacklisting subsequently manifest.

The rest of the paper is organized as follows. Section 2 contains the review of the literature. In Section 3 we discuss our data and present the identification strategy. Section 4 reports the empirical model and results. Section 5 concludes.

2. Related Literature

Blacklisting procedures have been introduced in 2000. Since that time, relatively few economic studies on the stigma effect have been produced.

The first theoretical and empirical discussion of the stigma effect as a controversial issue is found in Masciandaro (2005). The study highlight how in the aftermath of 9/11, growing attention has been paid to the role of lax financial regulation in facilitating money laundering and the financing of terrorism (criminal finance).

Two interacting principles are commonly described in the debate on the relationship between money laundering and regulation: a) illegal financial flows are facilitated by lax financial regulation; b) countries adopting lax financial regulation do not co-operate with the international effort aimed at combating criminal finance (International Monetary Fund 1998, Holder, 2003). These two principles characterize the mandate of the Financial Action Task Force (FATF) for the prevention of money laundering and terrorism finance.

On the one hand, in order to address the problems associated with criminal finance risks, it is fundamental to develop legal standards for regulation. FATF standards (Recommendations) have become the benchmark for measuring the degree of laxity of AML/CFT financial regulation in every single country setting.

On the other hand, faced with the problem of the lack of international harmonization and coordination, the FATF uses a list of specific criteria in order to monitor the compliance of countries with international standards. Those lists of compliance are commonly described as blacklists (Alexander, 2001; Masciandaro, 2005; and Verdugo Yepes, 2011). Blacklisting represents the cornerstone of the international regulation, with the effort to reduce the risk that some countries or territories can turn into havens for criminal financial activities. Blacklisting is based on the stigma effect, i.e. the threat for a listed country to face a drop in capital inflows and then the erosion of its competitive advantage after the inclusion in the list (Hampton and Christensen, 2002).

Here the possibility of the stigma paradox occurs. Focusing on the supply of regulation, the study notes that various jurisdictions, notwithstanding the blacklisting threat, delay or fail to change their financial rules, confirming their non-cooperative attitude (*reluctant friend effect*). Furthermore, although the fact that most jurisdictions in the blacklist enact regulatory measures in an effort to be removed from it, it remains to be proven that regulatory reforms are sufficient to guarantee a real change in the country non-cooperative attitude, with a decreasing appeal for black capital flows (*false friend effect*). The existence of these two consequences can nullify the stigma effect, producing stigma neutrality or even the stigma paradox.

The theoretical analysis under discussion develops the assumption that lax financial regulation may be a strategic dependent variable for national policymakers seeking to maximize the net benefits produced by such a policy, just like any other public policy choice. Therefore, given the structural features and endowments of their own country, certain policymakers may find it profitable to adopt financial regulations which accommodate the needs of opaque financial flows – whose existence is given by assumption – and therefore may choose to be a *de facto* BLC jurisdiction.

The potential incentives to be a BLC have been empirically tested using cross-section estimates, finding that the probability of being a BLC jurisdiction may be linked to specific country features (Masciandaro, 2005; Verdugo Yepes, 2011; Schwarz, 2011). The rationale for the strategy of being a BLC has been further explored from a theoretical point of view (Unger and Rawlings, 2008; Gnutzmann et al., 2010). Recently, the interactions between the FATF standards and national government activities have been further analyzed using a principal-agent framework (Ferwerda, 2012).

The economics of the stigma effect is analyzed in depth in Picard and Pieretti (2011), who focus on the incentives that banks located in a BLC have for complying with the AML/CFT regulation. The blacklisting practice is interpreted as an international pressure policy on the BLC bank and the stigma effect holds anytime the pressure policy is strong enough. More precisely, the stigma effect becomes effective when the reputational costs linked with the blacklisting procedures – which can harm banks' costumers – are higher than the costs of compliance. In the model, international policymakers act efficiently and thus implement optimal blacklisting pressure. In the real world, non-efficient policymakers are likely to exist, blacklisting pressure can be insufficient and the BLC might continue to attract financial flows, creating the stigma paradox.

The possibility of the stigma paradox has been empirically demonstrated in Rose and Spiegel (2006). Using bilateral and multilateral data from over 200 countries into a gravity framework, the study analyzes the determinants of the international capital flows, finding that a country status of tax haven and/or money launderer assigned by the international organizations can produce beneficial effects. The analysis confirms that the desire to circumvent national laws and regulations can be a driver in shifting financial assets abroad.

The search for the impact of the blacklisting has also been implemented in Kurdle (2008). Using ARIMA techniques on a sample of blacklisted countries, the study analyzes the financial effects of being listed and delisted. The results are inconclusive: the three effects – stigma effect, stigma paradox and stigma neutrality – can be equally found, depending on the time and the observed jurisdiction.

3. Data and Summary Statistics

3.1 Data sources

Our study compiles data from three main sources:

1. International banking statistics published by the Bank of International Settlements (BIS), which provides data regarding international banking flows for 126 countries (the list of the countries can be found in Appendix 1) for the period from 1996 to 2014.

The BIS Locational Banking Statistics publishes quarterly information on all balance sheet positions (and some off-balance sheet positions in the area of trustee business) which represent financial claims or liabilities vis-à-vis non-residents, as well as financial claims or foreign currency liabilities vis-à-vis residents. Banks report data to the central bank or monetary authority of the country where they are headquartered, and national aggregate data are then transmitted to the BIS.

The main dependent variable – Banking Flows – is constructed using data on Total Foreign Claims from the BIS consolidated banking statistics database (CBS). The relevant data are contained in Table 9A:S of the CBS (Total foreign claims, immediate borrower basis).

The bank flow variable is equal to the difference between the logarithm of total foreign claims in period $(t+1)$ and the logarithm of total foreign claims in period (t)

$$BankFlow_{i,t+1} = \log\left(\frac{TFC_{i,t+1}}{TFC_{i,t}}\right), \quad (1)$$

where $TFC_{i,t}$ is the value of Total Foreign Claims in year t for country *i*.

Foreign claims are financial claims on residents of countries other than the reporting country, i.e. claims on non-residents of the reporting country. In the CBS, foreign claims are calculated as the sum of cross-border and local claims (in all currencies) of reporting banks' foreign affiliates or, equivalently, of international claims and local claims denominated in local currencies.

2. Another data source we use in our analysis is Martin Čihák, Asli-Demirgüç-Kunt, Erik Feyen, and Ross Levine's "Global financial development Database" (GFDD). This is an extensive dataset of financial system characteristics for 205 countries from 1960 to 2010. The database enables to measure the size of financial institutions and markets, the degree to which individuals can and do use financial services, the efficiency of financial intermediaries and markets in intermediating resources and facilitating financial transactions, the stability of financial institutions and markets, and the level of macroeconomic stability.

3. Barth J.R., Caprio G. Jr., Levine R. "Banking Regulation Survey" 2000, 2003, 2007, 2012 provides data for the construction of the *Overall Activities Restrictions Index* and the *Independence of Supervisory Authority Index*.

The *Overall Activities Restrictions Index* is a dummy variable that takes value 1 if the banks of a country are subject to some restrictions in their operations, and zero if they are very free in their investment choices.

The *Independence of Supervisory Authority Index* expresses the extent to which the supervisory authority is independent from the government and legally protected from the banking industry.

In addition to those three main datasets, we also use a variety of other data sources. Specifically, we use the dataset constructed by Masciandaro et al. (2013) for banking sector supervision and regulation variables, World Bank and IMF data for macroeconomic controls and the dataset constructed by Mattes and Rodriguez (2014) to calculate cooperation variables.

3.2 Variables

3.2.1 International capital flows measure

The main goal of this research is to investigate the effect of soft regulation on the international capital flows. The international capital flows measure is used as our dependent variable. In the specific, we use growth rate of the total foreign claims obtained from the BIS database. The construction of the Bank Flow variable is described in equation (1).

Figure 1 (Appendix 1) represents the average value of International Bank Flows for all 126 countries in the period from 1996 until 2014.

Blacklisting

The variable of interest (our main predictor) is the dummy variable -- ***FATF*** -- that is equal to 1 if the country is listed in the Financial Action Task Force list of "Non-Cooperative Countries or Territories" and 0 if the country complies with FATF conditions. FATF variable is constructed using Financial Action Task Force reports "Review to Identify Non-Cooperative countries or

Territories: Increasing the Worldwide Effectiveness of Anti-Money Laundering Measures", published annually, in June, by the FATF. The report covers a period of 12 months, i.e. the report published in June 2009 represents the blacklisting status of the country during the last year, starting from June 2008 until June 2009. We assign the status of "listed" to the country in year(t), if in June of $(t+1)$ that country is in the FATF list.

The Financial Task Force List consists of countries perceived to be non-cooperative in the global fight against money laundering and terrorist financing. To evaluate the involvement of a country in terrorist financing and money laundering FATF has created a list of recommendations, which includes 40 recommendations on money laundering and 9 Special Recommendations on Terrorist Financing. The lack of cooperation by those countries manifests itself as unwillingness or inability to follow FATF recommendations.

According to the Forty recommendations on money laundering and Report on Non-cooperative Countries and Territories, countries are required to:

- Exclude the following loopholes in financial regulation:
  - Inadequate regulation and supervision of financial institutions;
  - Inadequate rules for the licensing and creation of financial institutions;
  - Inadequate customer identification requirements;
  - Excessive secrecy provisions regarding financial institutions;
  - Lack of efficient suspicious transaction reporting system;
- Respect the impediments set by other regulatory requirements in terms of:
  - Inadequate commercial law requirements for registration of business and legal entities;
  - Lack of identification of the beneficial owner(s) of legal and business entities;
- Remove the obstacles to the international co-operation
  - At administrative level;
  - At the judicial level (absence of criminalization of money laundering, laws and regulations prohibiting international exchange of information, presence of tax evasion);
- Avoid owning inadequate resources for preventing, detecting and repressing money-laundering activities.

The judicial system is a core element of detection for non-cooperative countries or territories. Countries that are characterized by the absence of laws regulating money laundering and tax evasion or preventing information sharing about suspicious business entities are more inclined to be included in the FATF Blacklist.

The Blacklisting of a country is not affected by the level of its capital flows. In this paper, we are interested in the effect of the Blacklisting on the growth rate of international capital flows. Since Blacklisting is an exogenous event for capital movements, the problem of endogeneity is not crucial in our analysis.

Descriptive statistics for the FATF variable and Bank Flows are represented in Table 1 (Appendix 2).

In total 45 countries have been listed since the introduction of the FATF List.

Seventeen of these countries are developed; the remaining 28 belong to the underdeveloped economies, according to the World Bank characterization. Other informative statistics about listing can be found in Appendix 1. It is worth noting that in 2006 there are no countries in the FATF list. According to the FATF, all countries were compliant with the requirements.

Figure 2 (Appendix 1) shows, for black listed countries only, the mean growth rate in bank flows relative to the years of blacklisting. This simple plot provides a first feel of the stigma effect and it uses a sort of monitoring period (two years before and after the blacklisting) to point it out. As a comparison, in red is drawn the average growth rate for all countries over the entire sample period.

3.2.2 Control Variables

- Financial Institution Depth, Efficiency, Stability

To describe the depth of the banking system of a country we use the measure from the GFDD - Bank Private Credit to GDP (BPC). BPC is equal to the amount of credit to private sector created by deposit money banks and other financial institutions normalized by GDP. It is calculated according to the following formula:

$$BPC_t = 0.5 * \frac{(PC_t/P_{et}) + (PC_{t-1}/P_{et-1})}{GDP_t/P_{at}}, \qquad (2)$$

where PC_t is credit to private sector in year t, P_{et} is end-of-period consumer price index (CPI), and P_{at} is average annual CPI.

We use Net Interest Margin (NIM) from the "Global Financial Development Database" as the measure of the financial institution efficiency. NIM is calculated as the bank's income that has been generated by non-interest related activities as a percentage of total income (net-interest income plus non-interest income). Non-interest related income includes net gains on trading and derivatives, net gains on other securities, net fees and commissions and other operating income.

As a measure of the stability of financial institutions, we use the Bank Z-Score (always from GFDD). It captures the probability of default of a country's banking system, calculated as a weighted average of the Z-scores of a country's individual banks (the weights are based on the individual banks' total assets). Z-score compares a bank's buffers (capitalization and returns) with the volatility of those returns.

- Supervision and Regulation of the Financial Sector

We use a set of variables from the four worldwide surveys conducted by Barth J.R., Caprio G. Jr., Levine R. "Banking Regulation Survey" in 2000, 2003, 2007, 2012. Data from those surveys are used to construct measures for various aspects of bank regulation and supervision, such as the *Overall Activities Restrictions Index* (ORA) and the *Independence of the Supervisory Authority Index* (ISA). Moreover, we build up an additional index that we called the *Supervisory Lightness Index* (SLI). That is an indicator, which takes values between 0 and 20, and it represents an inverse measure of the degree of lightness of the supervision system (Appendix 1 - Figure 5, provides detailed information about this index). As a measure of supervisory governance, we take hints from the paper of Quintyn et al. 2004.

3.2.3 Other country controls

We also include several country-level variables to control for differences in economic development.

We control for growth in the real per capita GDP, real interest rates and the real exchange rate as well as for the GDP deflator to take into account macroeconomic factors.

In addition, we use a dummy variable identifying the time-period after the introduction of the FATF Black List (dummy is equal to one for the periods after 2000 and 0 for the earlier period). Doing so we want to see whether the introduction of the Black list itself has had any impact on the international capital flows dynamic.

To further control for the banking and financial supervision regime, we use the FINFIU variable. That is a dummy variable that takes the value of 1 if a Financial Intelligence Unit exists and it has a financial nature (i.e. it is under an authority or an entity involved in financial markets discipline), and 0 otherwise.

To test the relevance of social and political parameters, we employ a set of six indicators of good governance that express quantitative measures for the political environment. They respectively indicate: government effectiveness; political stability; regulatory quality; voice and accountability; rule of law; and control of corruption but seen as a whole, they show the structural capacity of the government to formulate and implement sound policies. These indexes are extrapolated from the Worldwide Governance Index of the World Bank while the overall project of the WGI follows the proposal by Kaufmann et al. (2003).

Specific political factors are also included. Polity2 and durable represent two important proxies of the persistence and change in Political regimes. They are both extracted from Marshall et al. (2002).

Finally, an indicator for the size of the shadow economy is considered. Values raise from Schneider et al. (2010).

Descriptive statistics for control variables are presented in Table 2 (Appendix 2).

4. Empirical Analysis

4.1 Main Analysis

Here we empirically test the effect of the soft regulation (blacklisting) on the international bank flows. As a proxy for the blacklisting, we consider the abovementioned Financial Action Task Force Listing of countries. Countries not compliant with the requirements of the FATF are included in this list, while they can easily get out of it after fulfilling all the requirements.

For each of the considered years, we construct the Bank Flow Measure, which is equal to the logarithm of the fraction of Total Foreign Claims in year *t* over Total Foreign claims in year $(t+1)$. We use this measure as the dependent variable in the following regression:

$$Bank\ Flows_{i,t} = \alpha_0 + \alpha_1 FATF_{i,t} + \alpha_2 Regulation_{i,t} + \alpha_3 InstQuality_{i,t} + \gamma X_{i,t} + \varphi_i + \mu_t + \varepsilon_{i,t},$$

(3)

where i represents the country dimension of the sample, and t represents the time dimension;

✓ $FATF_{i,t}$ is a dummy variable equal to 1 if the country is listed in year *t*, and zero otherwise;

✓ $Regulation_{i,t}$ is a vector of regulation and supervision indexes, such as the Overall Activity Restrictiveness (OAR), the Independence of Supervisory Authority (ISA) and the Supervisory Lightness Index (SLI);

✓ $InstQuality_{i,t}$ is a vector of variables, representing features of the banking sector in country *i* at time *t*: Depth of Traditional Banking (Bank Private Credit to GDP, BPC), Degree of Innovativeness of Banking Activities (Net Interest Margin, NIM), Stability of the Banking Sector (Altman Z Score, Zscore);

✓ $X_{i,t}$ is a vector of independent variables representing other political and macroeconomic controls;

✓ φ_i is a country fixed effect and μ_t is the time-fixed effect (constructed by using dummy variables for each single year).

Table 3 (Appendix 2) lists estimated coefficients for regression (3).

In the Table 3 we provide results for three types of models: Random effect (1), model with year fixed effects (2), and model with year and country fixed effects (3). The main specification from which we draw our conclusions is the last one, with year and country fixed effects. This model helps us to capture all the unobservable country and year characteristics (for the details of our main regression, refer to Appendix 2 - Table 3, all the variants).

From the Table we see that the coefficient of the FATF variable is equal to 37.473 and it is statistically and economically significant. The coefficient is positive, meaning that the stigma effect does not appear. On the contrary, countries explore the regulatory arbitrage. Being listed can be considered as a sign of possibility to get additional profits by escaping taxation. In this case, the stigma paradox seems to be relevant.

It is interesting to note that the variables representing supervision are also significant. The coefficient for the Supervision Lightness Index is negative, while the coefficient of the Index of Independence of the Supervisory Authority is positive: the desire to take gains from a market friendly supervision managed by an authority independent from the local politicians seems to attract the capital flows.

At the same time, the traditional banking systems seem to be less attractive. As for the other controls, the variable Bank Private Credit has a negative coefficient. The larger the traditional banking system in the country, the slower the growth rate of International Banking Flows. The log of the GDP per capita is positively correlated with the Bank flow measure, while the presence of a relevant shadow economy slows down the growth in capital movements. On the one side the international capital flows seem to appreciate banking secrecy, but at the same time the bankers seem to dislike countries and territories where the significant role of the shadow economy can signalling weak rule of law, i.e. risks of instability and expropriation.

The next step of the analysis is to conduct robustness checks for our main results.

4.2 Robustness Check: Endogeneity of FATF measure

To proceed with the empirical analysis we should take into account several potential problems. In the previous part we highlight how, according to the official documents (recommendations), the fact that a country is included in the FATF Black List is not formally related to its international banking flows. In practice, however, we cannot exclude that the following can potentially be going on: FATF authorities track the same statistics as BIS. When an unusual growth in banking flows is noticed, the jurisdiction is flagged as a “priority for investigation”. This event automatically increases the probability of a country to be included in the Black List. Thus, an endogeneity problem can arise.

FATF determines the compliance of a country with the anti-money laundering and anti-terrorist financing regulation based on the already mentioned 49 Recommendations (last issue February 2013).

The decision to put a country into the Blacklist is based on a series of criteria. To prevent the event of blacklisting, a country should:

- Identify the risks, and develop policies and domestic coordination;
- Pursue money laundering, terrorist financing and the financing of proliferation;
- Apply preventive measures for the financial sector and other designated sectors;
- Establish powers and responsibilities for the competent authorities (e.g., investigative, law enforcement and supervisory authorities) and other institutional measures;
- Enhance the transparency and availability of beneficial ownership information of legal persons and arrangements;
- Facilitate international cooperation.

FATF authorities observe the performance of a country, including its international capital flows, and make a decision about blacklisting. Even though FATF policy can be considered exogenous to the internal policy of a country, we still may face endogeneity problems given those interconnections.

In order to address and fix these issues we apply the Instrumental Variables (IV) methodology and proceed through the empirical analysis. The first step is to propose an instrumental variable that is related as close as possible to the probability of being listed and that is uncorrelated with the international capital flows if not through the blacklisting. The second step, instead, uses the findings of the first stage in order to replicate the main regression, with the fitted values of the FATF variable being the main predictor.

4.2.1 Instrumental Variables analysis

FATF recommendations can be divided into several groups: AML/CFT policies and coordination; money laundering and confiscation; terrorist financing and financing proliferation; preventive measures; transparency and beneficial ownership of legal persons and arrangements; powers and responsibilities of competent authorities and other institutional measures; international cooperation. Based on this classification, we can conclude that regulatory factors and political variables determine the blacklisting status of the country.

One of the main factors that is reflected in the FATF recommendations is the level of International Cooperation that a country is involved in.

Goldstein (1992) has proposed the most used measure of international cooperation. In the paper, the author proposes a scale to classify countries based on their international relations. The conflict-cooperation scale has been created for the WEIS events data – World Event/Interaction Survey (1966-1978). The Goldstein scale incorporates both formal and informal cooperation and overall does well at capturing cooperation intensity. Recently the Goldstein scale has been applied to 10 million International Dyadic Events data (Bond, Bond, Oh, Jenkins and Taylor, 2003; King and Lowe, 2003) by Mattes and Rodriguez (2014). These data have been generated by the VRA Reader software, which classifies events found in Reuters Business Briefing news stories based on a typology called Integrated Data for event Analysis.

Mattes and Rodriguez have created two measures of international cooperation for the period to 1990 to 2004: the Goldstein Scale for pair of countries (GS_{ijt}) and a Dummy variable that is equal to 1 if countries cooperate in year t and 0 otherwise $(DummyInternCoop_{ijt})$.

For our purposes, we construct two indexes using variables from Mattes and Rodriguez.

The first index is set up taking information from the dummy variable. The international cooperation index for the country i in year t $(IC1_{it})$ is equal to:

$$IC1_{it} = \frac{\sum_j DummyInternCoop_{ijt}}{Total\ \#of\ countries}, \tag{3}$$

where j identifies the counterparty country for the country i.

The second index is similar in construction but more precise, since it takes into account the value of the Goldstein scale for each pair of cooperating countries. The International Cooperation Goldstein index for the country i in year t $(IC2_{it})$ is equal to:

$$IC2_{it} = \frac{\sum_j GS_{ijt} * DummyInternCoop_{ijt}}{Total\ \#\ of\ countries} \tag{4}$$

These indexes are two of the measures that can be used as IVs for the FATF listing event and, notwithstanding their limited availability, we strongly rely on them in this step of the analysis.

Another factor that affects the probability of being blacklisted is the political influence of the country.

The UN Security Council “takes the lead in determining the existence of a threat to peace or act of aggression”.

Being a representative state in the UN Security Council gives a country the right to imply a veto. Thus, being a member of the UN Security Council gives a country the power, among the others, to inflict anti-terrorist policies to other countries. We define the membership of the country i in the UN Security Council in year t as $UNSCmem_{it}$, a dummy variable, which is equal to 1 if the country i is actually a member and 0 otherwise. The $UNSCmem$ forms our third instrumental variable.

In the same way as in the case of the UN membership, being mentioned on the agenda of the UN Security Council can affect the probability of being blacklisted. The dummy variable $UNSCagenda_{it}$ is equal to 1 if the country i is on the agenda of the UN Security Council meeting in year t and it is equal to 0 otherwise. The $UNSCagenda$ completes our set of four instrumental variables.

To summarize, the IVs used to instrument the probability of being included in the FATF Black List are as in the Table 4 (Appendix 2).

The IV analysis is performed using a two-stage least squares (2SLS) regression.

The first stage is used to build up our new variable, $\widehat{FATF}$, that comes as a linear combination of the instrumental variables. In the specific, we generate four different predictions of $\widehat{FATF}$: the first one takes into account only the variables related to the UN Security Council; the second adds to the predictors the Goldstein scale; the third substitutes the Goldstein scale with the dummy cooperation; finally, in the fourth column we set up a linear combination of all of our four instruments. Results related to the first stage are presented in Table 5.

Those estimates are generated by means of OLS. In order to provide a further degree of consistency to our analyses, Appendix 2 reports the alternative specification that uses logit rather than the OLS (Table 5.a).

Once obtained our instrument $\widehat{FATF}$ in all its different forms, the second stage consists in regressing the Bank flows measure on them. Our main equation becomes:

$$Bank\ Flows_{i,t} = \alpha_0 + \alpha_1 \widehat{FATF}_{it} + \alpha_2 Regulation_{i,t} + \alpha_3 InstQuality_{i,t} + \gamma X_{i,t} + \varphi_i + \mu_t + \varepsilon_{i,t}, \quad (5)$$

Where the original FATF variable has been substituted by the new $\widehat{FATF}$, while all the other controls remain the same.

Results for the second stage are presented in Table 6 (Appendix 2). As before, Appendix 2 - Table 6.a reports the outcomes when operating logit in the first stage. In that case, a further correction for biased standard errors has been implemented (Wooldridge, 2002).

The coefficient of the $\widehat{FATF}$ variable that is associated to the International capital flows is always positive and never statistically significant. It is quite difficult to make inference in favour of the stigma paradox and/or to the stigma neutrality. However, we have strong evidence against the stigma effect. Our findings, those of Table 3 and those of Table 6, seem to suggest that, ultimately, a stigma effect for the black listed countries does not exist.

To give more consistency to our findings, at the end of the tables we provide some statistics about the IV regressions.

The under-identification test is an LM test of whether the equation is identified, i.e., that the excluded instruments are "relevant", meaning correlated with the endogenous regressors. The test is essentially the test of the rank of a matrix: under the null hypothesis that the equation is under-identified. Under the null, the statistic is distributed as chi-squared with degrees of freedom equal to $(L1 - K1 + 1)$. A rejection of the null indicates that the matrix is full column rank, i.e., the model is identified.

The Sargan-Hansen test is a test of over-identifying restrictions. The joint null hypothesis is that the instruments are valid instruments, i.e., uncorrelated with the error term, and that the excluded instruments are correctly excluded from the estimated equation. Under the null, the test statistic is distributed as chi-squared in the number of $(L - K)$ over-identifying restrictions. A rejection casts doubt on the validity of the instruments. In our case, the null is never rejected.

Finally, the endogeneity test, which estimates are reported in a separate table. Under the null hypothesis that the specified endogenous regressors can actually be treated as exogenous, the test statistic is distributed as chi-squared with degrees of freedom equal to the number of regressors tested. As we can see, the hypothesis of the exogeneity of the FATF variable can never be rejected, providing additional validity to our primordial analysis.

4.3 Robustness check: International Banking Inflows and Outflows

The measure of international bank flows is equal to the growth rate of a country's total foreign claims. Total foreign claims comprise cross-border claims plus local claims in local and foreign currencies. Cross-border claims are asset positions vis-à-vis banks and non-banks located in a country other than the country of residence of the reporting banking office; they are also referred to as "external" assets. Local claims include Asset positions with a counterparty (bank or non-bank) located in the same country as the banking office, denominated in local and foreign currency.

We can disentangle two effects of the FATF Listing. The first regards the bank outflows, the second is related to the bank inflows.

With regard to the former, the blacklisting effect can be further divided. On the one hand, the listing procedure directly influences the external bank outflows, i.e. the growth rate of External Assets. On the other hand, even the internal flows between domestic and foreign banks react to the blacklisting of the host country. The effects, as we can see in a moment, are different.

Having those definitions in mind, to ascertain the effect of blacklisting on outflows and inflows separately, we construct two variables.

The first variable, Bank External Outflow, is equal to the difference between the log of external assets (all sectors) in period $(t+1)$ and logarithm of external assets (all sectors) in period (t):

$$ExtOutflow_{i,t} = \log\left(\frac{ExternalAssets_{i,t}}{ExternalAssets_{i,t-1}}\right), \tag{8}$$

where $ExternalAssets_{i,t}$ is the value of External Assets (all sectors) in year t for the country *i*.

Empirical results considering External Outflows as the dependent variable of our main regression are presented in Table 7 (Appendix 2).

As we can see, the FATF coefficient remains positive. Moreover, FATF Listing does not significantly affect External Outflows.

Since Banking outflows are equal to the sum of External Outflows and Internal outflows and we do not observe any particular effect of FATF blacklisting on External Outflows, FATF Blacklisting should strongly affect Internal Banking Outflows. This result can be interpreted as reflecting the increase in the value and number of transactions between domestic and foreign local banks after the listing event.

The main reason for blacklisting are tax evasion and money laundering. When a country is included into the FATF List, domestic banks seem to decrease their international transactions. However, that does not mean that the country automatically becomes compliant. It means that banks transfer funds through subsidiaries of foreign banks located in the country. After that, such subsidiaries transfer money to the bank's headquarters. In that way, transactions stay undetected, the country can be excluded from the FATF List and banks can continue their illegal transactions.

We can say that FATF Listing is significant in affecting Internal Banking Outflow and does not affect External Banking Outflow. That is just a simple hypothesis.

Bank Inflows represent an alternative measure of International Bank Flows. To measure Bank Inflows we can use the External Liabilities data from the BIS Locational Banking Statistics. As regards the analytical formula, it perfectly reflects that of bank outflows.

Model estimation results when considering Bank Inflows as the dependent variable are presented in Table 8.

According to the results, FATF Listing has not a significant impact on the External Inflows as well.

4.4 Robustness check: Bank flows lags

Our final robustness check takes into account the lag values of international banking flows.

The idea is that to investigate the possibility that lagged growth in bank inflows predicts blacklisting, and that if growth in bank inflows is more or less constant for the affected countries, the stigma effect could simply capture this persistence.

Results of our main regression when considering also the lagged values of bank flows, are presented in Appendix 2 - Table 9.a and subsequent.

Our FATF measure appears to be often significant in explaining Bank flows. Moreover, when considering our full specification, with all the control variables, its coefficient turns out to be positive. Once more we cast doubts on the existence of the stigma effect.

5. Conclusion

This study analysed cross-border capital flows in order to verify the existence and direction of the effect of the soft regulation promoted by international organizations against banking secrecy that characterized the so called tax and financial heavens. This effect is called in the literature Stigma Effect.

Banking secrecy can be considered a hot and controversial topic for the international debate in the last decades. Particularly in the aftermath of the Global Financial Crisis the advanced economies declared that the fight against bank secrecy and offshore countries and territories is a political priority and the use of the blacklisting its main tool. In addition, recently, after the publication of the so-called Panama Papers, the German and French governments pressed OECD nations to pull together a common blacklist of countries that breach the international transparency standards. However, is the blacklisting effective?

We studied the relationship between international capital movements and FATF (Financial Action Task Force) listing/delisting events in 126 countries in the years from 1996 to 2014. We test whether international banking activities respond to the "name and shame" approach, which has been introduced to combat money laundering and terrorism finance. Consequently, we wonder if it is possible to detect financial gains for a country that implements AML/CFT policies consistent with the international standards. To understand the effects that FATF decisions produce over listed countries, we focus on how reacted to higher potential costs that can emerge (disappear) when a country is listed (delisted). Our empirical conclusion is that in general the stigma effect does not exist.

Therefore, is the era of banking secrecy definitively over, as a G20 official document stated in 2009? Probably not. The economic rationale of our results is easy to found out: if it is assumed that banking secrecy is the result of market mechanisms, it is easy to forecast that the worldwide demand and supply of banking secrecy are likely to be relevant for a long time to come, and consequently the existence of tax and financial heavens.

The bottom line is that the growth of criminal and illegal activities systematically generates the demand of banking secrecy, while economic and political incentives can motivate national politicians and international banks to demand and to supply banking secrecy. Using a soft law tool as the blacklisting is likely to be a weak policy solution for a structural problem with deep roots in the incentive structures of both offshore and onshore countries. In fact, banking secrecy is unlikely to disappear; it is truer to describe it as a dynamic variable with its booms and busts motivated by the changing preferences of national policymakers. Banking secrecy is a like a tango: it takes two to dance it. In addition, the blacklisting is unlikely to stop the music.

6. References

- Alexander, K. (2001), The International Anti-Money Laundering Regime: The Role of the Financial Action Task Force, *Journal of Money Laundering Control*, vol.4, n.3, 231-248.

- Barth J.R., Caprio J. and R. Levine (2006), Rethinking Bank Regulation: Till Angels Govern, Cambridge, Cambridge University Press.

- Bond D., Bond J., Oh C., Jenkins J.C. and Taylor C.L. (2003), Integrated data for events analysis (IDEA): An event typology for automated events data development, *Journal of Peace Research*, *40*(6), 733-745.

- Brummer C. (2012), *Soft Law and the Global Financial System: Rule Making in the 21st Century*, Cambridge University Press, New York.

- Chitu l., Eichengreen B. and Mehl A.J. (2013), History, Gravity and International Finance, NBER Working Paper Series, National Bureau of Economic Research, n. 18697.

- Das M.U.S., Quintyn M.M. and Chenard M.K. (2004), *Does regulatory governance matter for financial system stability? An empirical analysis* (No. 4-89), International Monetary Fund.

- FATF (2000), Report on Non-Cooperative Countries and Territories, Financial action Task Force, FATF/OECD, Paris.

- FATF (2012), *International Standard on Combating Money Laundering and the Financing of Terrorism and Proliferation. The FATF Recommendations*, Financial Action Task Force, FATF/OECD, Paris.

- Ferwerda J. (2012), The International Fight Against Money Laundering, in The Multidisciplinary Economics of Money Laundering, Chapter 7, Dissertation Series, Tjalling C. Koopmans Institute, School of Economics, Utrecht University, Ridderprint, Ridderkerk, 97-118.

- Goldstein J.S. (1992), A conflict-cooperation scale for WEIS events data, .*Journal of Conflict Resolution*, *36*(2), 369-385.

- Greene E.F. and J.L. Boehm (2012), The Limits of "Name- and- Shame" in the International Financial Regulation, *Cornell Law Review*, 97(5), 1083-1140.

- Gnutzmann H., K.J. Mc Carthy, B. Unger (2010), Dancing with the Devil: Country Size and the Incentive to Tolerate Money Laundering, *International Review of Law and Economics*, 30, 244-252.

- Holder, W.E (2003), The International Monetary Fund's Involvement in Combating Money Laundering and the Financing of Terrorism, *Journal of Money Laundering Control*, vol.6, n.4, 383-387.

- Hampton M.P and J. Christensen (2002), Offshore Pariahs? Small Island Economies, Tax Havens, and the Reconfiguration of Global Finance, *World Development*, 30(6), 1657-1673.

- Houston J.F., Lin C. and Y. Ma (2011), Regulatory Arbitrage and International Bank Flows, *Journal of Finance*, forthcomings.

- IMF (1998), *Money Laundering. The Importance of International Countermeasures*, address by Michel Camdessus, Plenary Meeting of the FATF, International Monetary Fund, Washington D.C., pp. 1-4.

- Kaufmann D., Kraay A. and M. Mastrucci, (2008), Governance Matters VII: Aggregate and Individual Governance Indicators, 1996-2007, Policy *Research Working Paper Series,* World Bank, Washington, DC, n. 4654.

- King, G. and Lowe W. (2003), An automated information extraction tool for international conflict data with performance as good as human coders: A rare events evaluation design, *International Organization*, *57*(03), 617-642.

- KPMG, (2011), *Global Anti – Money Laundering Survey*, at kpmg.com.

- Kudrle, R., (2009), Did Blacklisting Hurt the Tax Havens?, *Journal of Money Laundering Control*, Vol. 12 (1), 33-49.

- Marshall M.G. and Jaggers K. (2002), Polity IV project: Political regime characteristics and transitions, 1800-2002.

- Masciandaro D., (2005), False and Reluctant Friends? National Money Laundering Regulation, International Compliance and Non-Cooperative Countries, *European Journal of Law and Economics*, 2005, n.20, 17-30.

- Masciandaro, D., (2008), Offshore Financial Centres: the Political Economy of Regulation, *European Journal of Law and Economics,* Vol. 26, 307-340.

- Masciandaro D., (2013), Is the Anti Money Laundering Compliance Convenient?, *IDB Discussion Paper Series*, Inter-American Development Bank, Washington D.C., n.311 2013.

- Masciandaro D, Takats E. and Unger B., (2007), *Black Finance. The Economics of Money Laundering,* Edward Elgar, Cheltenham.

- Masciandaro D., Pansini R.V. and M. Quintyn (2013), The Economic Crisis: Did Supervisory Architecture and Governance Matter, *Journal of Financial Stability*, forthcoming.

- Mattes M. and Rodríguez M. (2014), Autocracies and international cooperation, *International Studies Quarterly*, *58*(3), 527-538.

- Milesi Ferretti G.M and Tille C., (2011), The Great Retrenchment: International Capital Flow During the Global Financial Crisis, Working Paper Series, Hong Kong Institute for Monetary Research, n.38.

- Papaioannou E. (2009), What Drives International Financial Flows? Politics, Institutions and Other Determinants, *Journal of Development Economics*, 88, 269-281.

- Picard P.M and P. Pieretti (2011), Bank Secrecy, Illicit Money and Offshore Financial Centres, *Journal of Public Economics*, 95 (7-8), 942-955.

- Powell J.H. (2013), Anti-Money Laundering and the Banking Secrecy Act, Board of Governors of the Federal Reserve System, Committee on Banking, Housing and Urban Affairs, U.S. Senate, Washington D.C., March 7, mimeo.

- Qureshi M.S., Ostry J.D., Ghosh A.R. and M. Chamon (2011), Managing Capital Inflows: The Role of Capital Controls and Prudential Policies, *NBER Working Paper Series*, n.17363.

- Reinhardt D., Ricci L.A. and T. Tressel (2010), International Capital Flows and Development: Financial Openness Matters, *IMF Working Paper Series*, International Monetary Fund, n.235.

- Rose A.K. and M. Spiegel (2006), Offshore Financial Centers: Parasites or Symbionts?, *Economic Journal,* 117(523), 1310-1335.

- Schneider F., Buehn A. and Montenegro C.E. (2010), Shadow Economies all over the World: New Estimates for 162 Countries from 1999 to 2007, *World Bank Policy Research Working Paper Series, Vol*.

- Schwarz P. (2011), Money Launderers and Tax Havens: Two Sides of the Same Coin?, *International Review of Law and Economics*, 31, 37 – 47.

- Unger B. and Rawlings G. (2008), Competing for Criminal Money, *Global Business and Economics Review*, 10(3), 331-352.

- Verdugo Yepes C. (2011), Compliance with the AML/CFT International Standard: Lessons from a Cross-Country Analysis, *IMF Working Paper Series*, International Monetary Fund, n.177.

- Wooldridge J.M. (2002), Econometric analysis of cross section and panel data, MIT press.

Appendix 1

List of countries in the sample

1.	Argentine	**33.**	Cyprus	**65.**	Jordan	**97.**	Peru
2.	Aruba	**34.**	Czech Republic	**66.**	Kazakhstan	**98.**	Portugal
3.	Albania	**35.**	Denmark	**67.**	Kenya	**99.**	Romania
4.	Algeria	**36.**	Dominica	**68.**	Korea, South	**100.**	Russia
5.	Angola	**37.**	Egypt	**69.**	Kuwait	**101.**	Saudi Arabia
6.	Australia	**38.**	Dominican Republic	**70.**	Kyrgyzstan	**102.**	Seychelles
7.	Austria	**39.**	El Salvador	**71.**	Latvia	**103.**	Singapore
8.	Azerbaijan	**40.**	Estonia	**72.**	Lebanon	**104.**	Slovakia
9.	Bahamas	**41.**	Ecuador	**73.**	Lesotho	**105.**	Slovenia
10.	Bahrain	**42.**	Finland	**74.**	Liechtenstein	**106.**	South Africa
11.	Bangladesh	**43.**	France	**75.**	Lithuania	**107.**	Spain
12.	Barbados	**44.**	Germany	**76.**	Luxembourg	**108.**	Sri Lanka
13.	Belarus	**45.**	Ghana	**77.**	Macau	**109.**	St. Lucia
14.	Belgium	**46.**	Gibraltar	**78.**	Macedonia	**110.**	Sweden
15.	Belize	**47.**	Greece	**79.**	Malaysia	**111.**	St. Vincent and the Grenadines
16.	Bhutan	**48.**	Grenada	**80.**	Malta	**112.**	Switzerland
17.	Bermuda	**49.**	Guernsey	**81.**	Mauritius	**113.**	Surinam
18.	Bolivia	**50.**	Guatemala	**82.**	Mexico	**114.**	Tajikistan
19.	Bosnia and Herzegovina	**51.**	Guyana	**83.**	Moldova	**115.**	Thailand
20.	Botswana	**52.**	Hungary	**84.**	Morocco	**116.**	Trinidad and Tobago
21.	Brazil	**53.**	Haiti	**85.**	Namibia	**117.**	Turkey
22.	Bulgaria	**54.**	Iceland	**86.**	Nepal	**118.**	Turkmenistan
23.	Burundi	**55.**	Honduras	**87.**	Netherlands	**119.**	Ukraine
24.	Cambodia	**56.**	India	**88.**	New Zealand	**120.**	United Kingdom
25.	Canada	**57.**	Indonesia	**89.**	Nicaragua	**121.**	United States
26.	Cayman Islands	**58.**	Ireland	**90.**	Nigeria	**122.**	Uruguay
27.	China	**59.**	Isle of Man	**91.**	Oman	**123.**	Vanuatu
28.	Chile	**60.**	Israel	**92.**	Pakistan	**124.**	Venezuela
29.	Colombia	**61.**	Italy	**93.**	Panama	**125.**	Vietnam
30.	Croatia	**62.**	Jamaica	**94.**	Philippines	**126.**	Zimbabwe
31.	Costa Rica	**63.**	Japan	**95.**	Paraguay		
32.	Cuba	**64.**	Jersey	**96.**	Poland		

Figure 1. Average value of International Bank Flow, 1996-2014

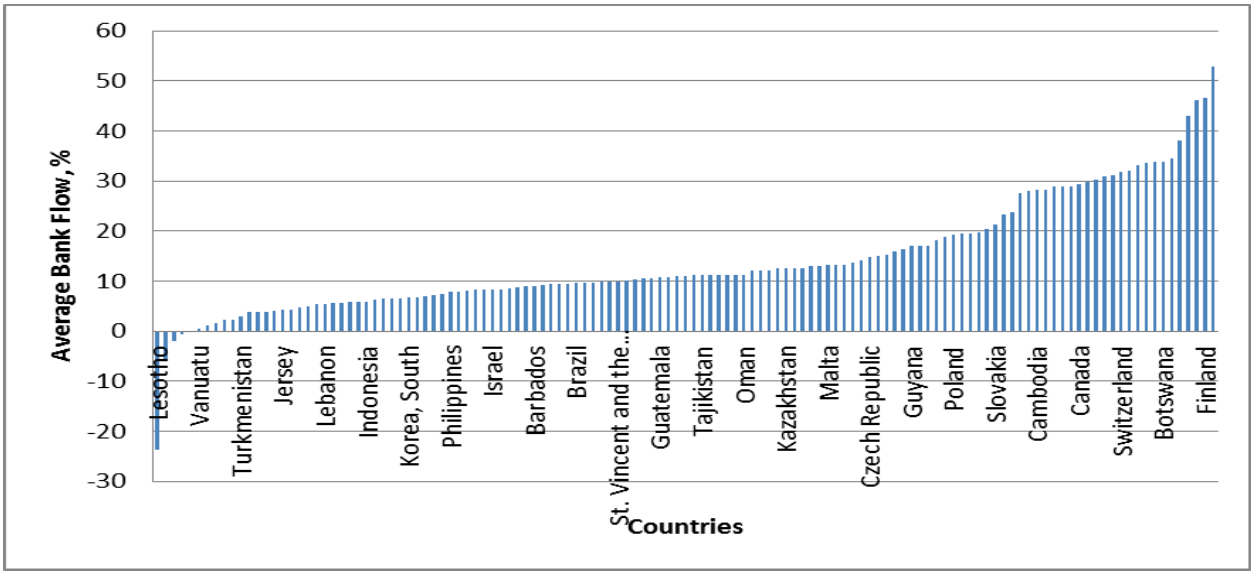


Figure 2. Mean growth rate of International Bank Flows, black listed countries in the years of blacklisting

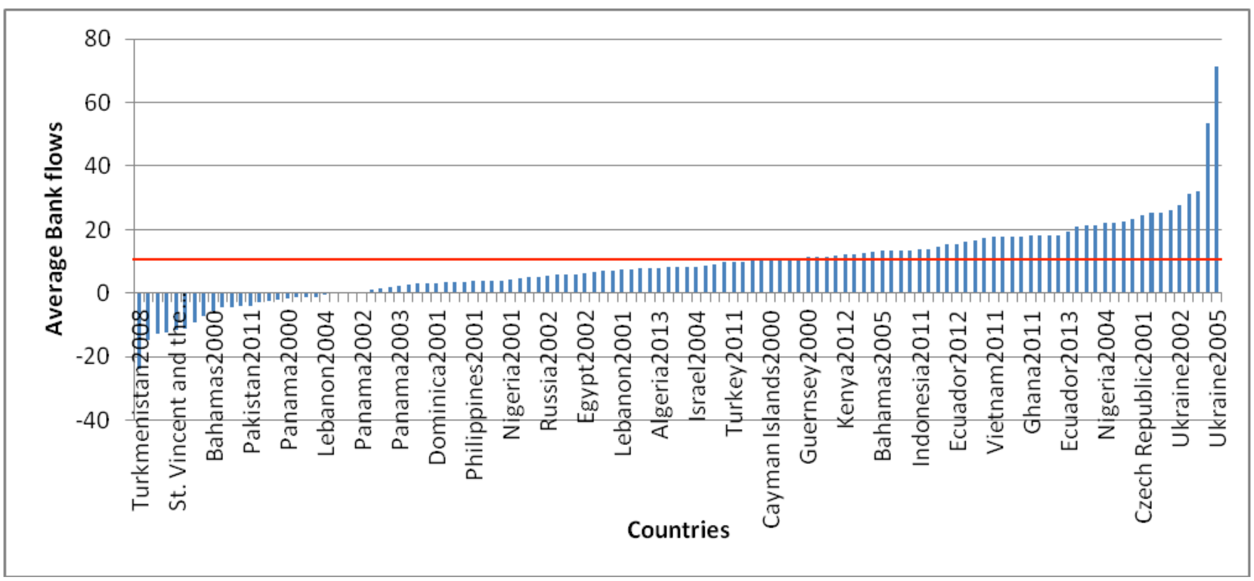


Figure 3. Statistics for listed and non-listed countries

Characteristics	Number of countries	OECD Countries	Emerging markets
Listed	45	6	11
One year listed	19	3	3
Two years listed	6	1	1
More than two years listed	20	2	7

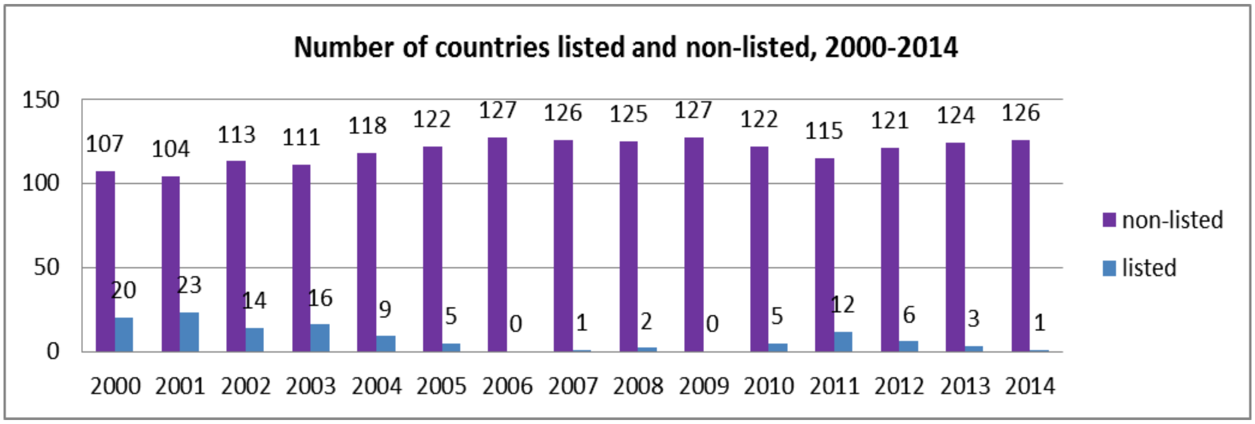


Figure 4. Bank Flow Measure: listed and non-listed countries

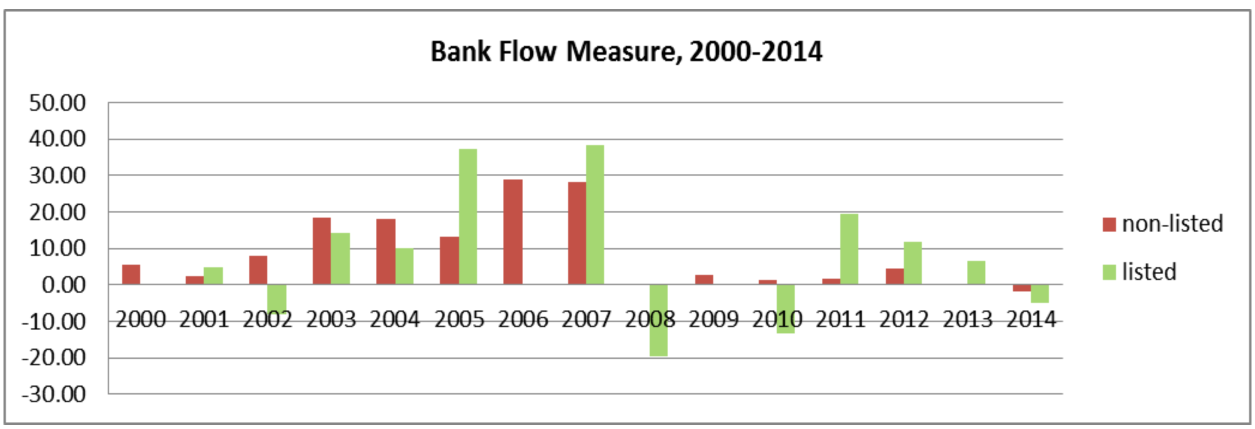


Figure 5. Some Independent Variables: listed and non-listed countries

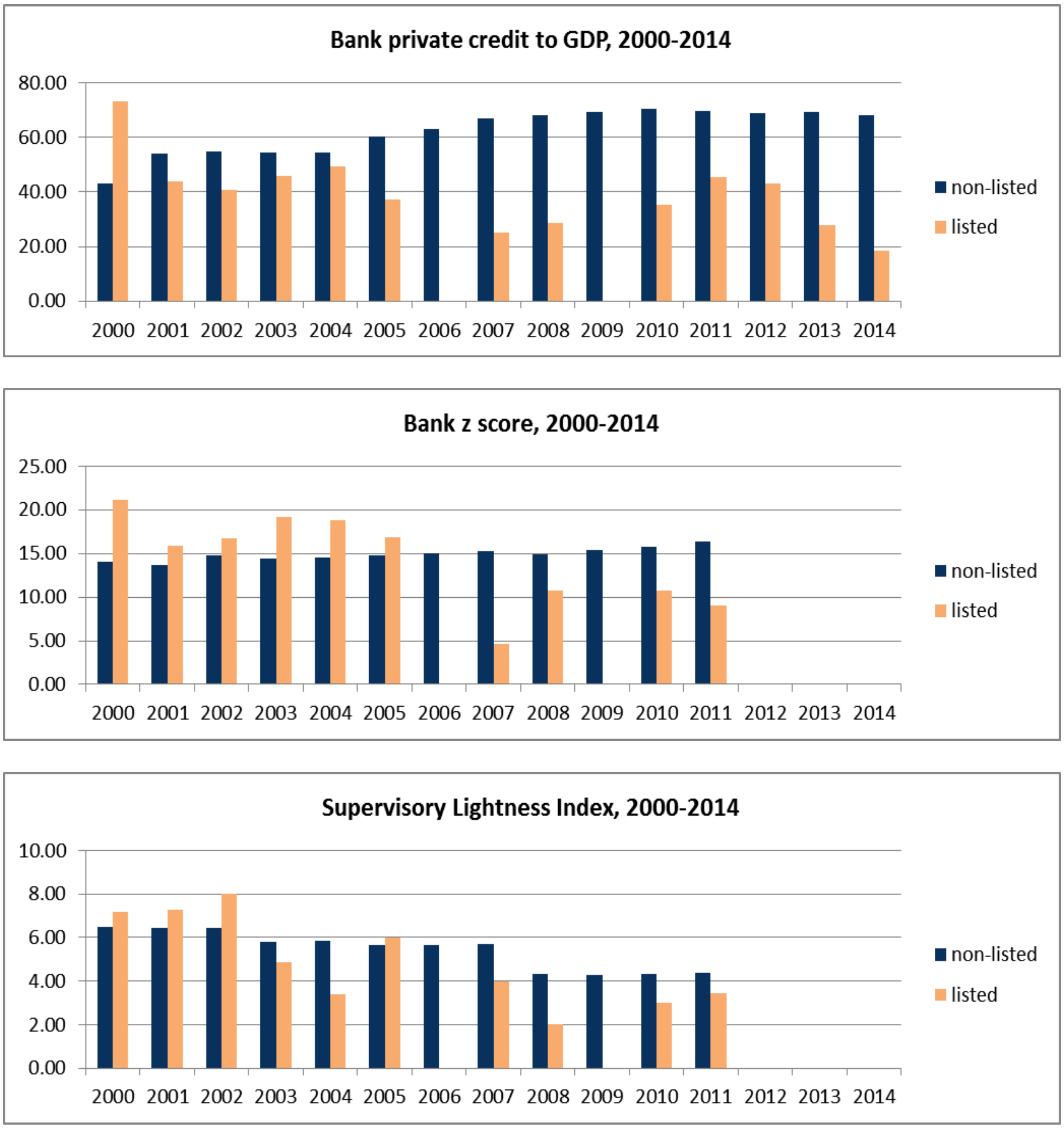
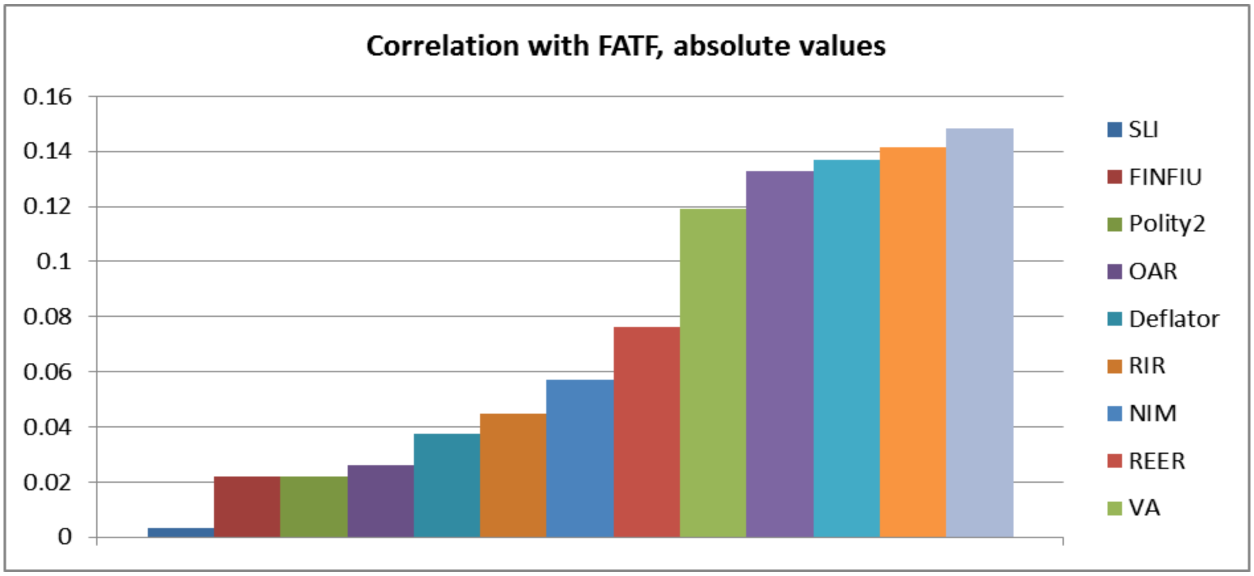


Figure 6. Correlation of FATF Listing variable with other independent variables


Correlation with FATF, absolute values
0.16
0.14
0.12
0.1
0.08
0.06
0.04
0.02
0
SLI
FINFIU
Polity2
OAR
Deflator
RIR
NIM
REER
VA

Appendix 2

Table 1. Descriptive Statistics: Bank Flows and FATF

Variable		Mean	Std. Dev	Min	Max	Observations
Bankflow	overall	14.48147	57.10873	-192.369	962.2409	N = 2334
	between		12.68277	-28.2781	58.97776	n = 126
	within		55.76565	-181.94	917.7446	T-bar = 18.5238
FATF	overall	0.048872	0.215646	0	1	N = 2394
	between		0.085563	0	0.368421	n = 126
	within		0.198084	-0.31955	0.996241	T = 19

Table 2. Control variables, descriptive statistics

Variable		Mean	Std.Dev.	Min	Max	Observations
OAR	overall	0.7691	0.421528	0	1	N = 1767
	between		0.267624	0	1	n = 121
	within		0.325021	-0.0434	1.5816	T = 14.6033
ISA	overall	1.82147	0.820551	0	3	N = 1714
	between		0.609041	0.363636	3	n = 121
	within		0.567102	-0.0452	3.954804	T = 14.1653
SLI	overall	5.838729	2.609646	0	14	N = 1699
	between		1.764543	2.307692	10.3125	n = 121
	within		1.942521	-1.22377	12.15123	T = 14.0413
BPC	overall	58.81332	46.31515	0.557351	311.063	N = 2143
	between		42.84895	1.478462	203.6033	n = 118
	within		17.92076	-50.2838	231.0052	T-bar = 18.161
NIM	overall	4.38293	2.94119	0.006614	23.31961	N = 1513
	between		2.573156	0.587567	13.15254	n = 121
	within		1.5104	-8.73734	15.14527	T = 12.5041
BZS	overall	15.01105	10.44965	-11.5741	58.71029	N = 1516
	between		9.697505	-1.61389	48.28198	n = 120
	within		3.787908	-0.0264	43.25155	T = 12.6333
GDPgro	overall	8.606893	1.478962	4.234848	11.96868	N = 2289
	between		1.407985	5.124587	11.53326	n = 122
	within		0.477251	5.69425	10.17332	T = 18.7623
RIR	overall	7.455217	22.16294	-94.2199	572.9363	N = 1864
	between		16.07804	-14.1652	164.2393	n = 112
	within		17.14335	-133.616	416.1522	T-bar = 16.6429
REER	overall	102.4263	13.83995	42.40215	179.9582	N = 1379
	between		6.352384	74.00537	128.7365	n = 126
	within		11.16037	58.99066	161.6395	T = 10.9444
GDPdef	overall	11.37652	118.9026	-27.6318	5399.507	N = 2270
	between		34.5289	-0.90923	368.516	n = 122

	within		113.7147	-364.558	5042.367	T-bar = 18.6066
FINFIU	overall	0.326351	0.468976	0	1	N = 2387
	between		0.381941	0	1	n = 126
	within		0.27394	-0.62102	1.221088	T = 18.9444
VA	overall	0.235476	0.930759	-2.21028	1.826408	N = 1950
	between		0.92052	-1.9644	1.614059	n = 123
	within		0.162053	-0.60773	1.003748	T = 15.8537
PS	overall	0.080456	0.920816	-2.81208	1.665204	N = 1950
	between		0.891346	-2.07404	1.480144	n = 123
	within		0.262076	-1.33711	1.48025	T = 15.8537
GE	overall	0.310465	0.944976	-1.72685	2.429651	N = 1947
	between		0.935112	-1.46447	2.158809	n = 123
	within		0.162076	-0.48395	1.194526	T = 15.8293
RQ	overall	0.308625	0.916619	-2.21026	2.247345	N = 1947
	between		0.899078	-2.00456	1.917077	n = 123
	within		0.190098	-0.4191	1.461987	T = 15.8293
RL	overall	0.20086	0.983393	-1.9094	2.001923	N = 1950
	between		0.979434	-1.5788	1.938847	n = 123
	within		0.153505	-0.49508	1.069367	T = 15.8537
CC	overall	0.236149	1.026974	-1.81587	2.585612	N = 1947
	between		1.012835	-1.32798	2.444269	n = 123
	within		0.186644	-0.64945	1.335162	T = 15.8293
Polity2	overall	5.235146	6.033124	-10	10	N = 1986
	between		5.765381	-10	10	n = 105
	within		1.819977	-7.02801	14.44567	T-bar = 18.9143
Durable	overall	31.14647	34.99643	0	205	N = 2014
	between		34.27294	0	196	n = 106
	within		7.785832	-14.1693	82.83069	T = 19
ShadowEc	overall	30.89012	13.32441	8.1	67.7	N = 972
	between		13.36793	8.544444	66.06667	n = 109
	within		1.143708	24.85679	40.23457	T = 8.91743

Table 3. Estimation results: Main Model

VARIABLES	**RE (1)** BankFlowMeasure	**Year FE (2)** BankFlowMeasure	**Year and Country FE (3)** BankFlowMeasure
FATF	11.6621*	26.0158***	37.4726***
	(6.086)	(7.791)	(12.537)
OAR	2.7667	5.8246	7.2191
	(7.520)	(7.529)	(8.608)
ISA	1.0617	2.8250	11.5833*
	(3.603)	(3.631)	(6.612)
SLI	-0.9130	-1.0171	-3.5938**
	(1.230)	(1.217)	(1.504)
BPC	-0.0985	-0.1253	-1.8172**
	(0.121)	(0.117)	(0.727)
NIM	-0.3611	-0.1438	-0.2752
	(0.793)	(1.129)	(1.430)
BZS	0.2150	0.1366	-0.8322
	(0.208)	(0.219)	(0.755)
Log GDP per Capita	-0.5727	3.7974	74.4295**
	(4.136)	(4.247)	(32.545)
Real Interest rate	-0.5917***	-1.0151***	-0.2746
	(0.209)	(0.325)	(0.833)
Real Exchange Rate	0.4181**	0.2183	0.0807
	(0.207)	(0.262)	(0.502)
GDP deflator	-0.2442	-0.8016**	-0.2973
	(0.252)	(0.403)	(0.583)
FINFIU	1.1201	8.8133*	11.4278
	(5.303)	(5.303)	(19.295)
Voice & Accountability	4.0202	2.7320	0.2718
	(12.022)	(13.475)	(24.094)
Political Stability	14.1852**	13.9066**	26.9338
	(6.297)	(6.339)	(29.186)
Government Effectiveness	6.2568	6.9144	28.9545
	(14.665)	(14.806)	(33.622)
Regulatory Quality	12.3506	1.4325	0.3891
	(14.513)	(15.607)	(34.588)
Rule of Law	-42.7462**	-38.9196**	4.8749
	(17.990)	(17.917)	(34.933)
Control of Corruption	19.1584	19.0729	19.6466
	(16.071)	(15.122)	(29.555)
Polity2	-0.0368	0.1496	-0.5771
	(1.257)	(1.502)	(4.647)
Durable	0.3121***	0.3037***	-1.8905
	(0.101)	(0.082)	(2.396)
Size of Shadow Economy	-0.0633	-0.1244	-7.5073*
	(0.387)	(0.384)	(4.198)
Constant	-20.2945	-35.6012	-255.8468
	(44.738)	(50.486)	(284.460)
Observations	396	396	396
Number of ID	85	85	85
Country FE	NO	NO	YES
Year FE	NO	YES	YES
R-squared (overall)	0.102	0.206	0.290

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Table 3.a. Capital flows: Random effects

VARIABLES	(1) BankFlowMeasure	(2) BankFlowMeasure	(3) BankFlowMeasure	(4) BankFlowMeasure	(5) BankFlowMeasure	(6) BankFlowMeasure	(7) BankFlowMeasure
FATF	-15.9081***	-7.9442	-6.2048**	-2.7330	5.0075	10.2584	11.6621*
	(3.283)	(5.075)	(3.017)	(5.872)	(7.777)	(6.737)	(6.086)
OAR	-0.4706	-9.2688	5.0483	-5.5962	10.1305*	0.5581	2.7667
	(4.907)	(8.587)	(3.854)	(7.221)	(5.573)	(8.236)	(7.520)
ISA	2.5036	7.7546*	3.8051*	5.4351	1.7542	1.6227	1.0617
	(1.998)	(4.564)	(1.961)	(4.355)	(3.114)	(4.023)	(3.603)
SLI	1.4012**	0.2225	0.5907	-0.1942	0.5391	-1.4437	-0.9130
	(0.562)	(1.814)	(0.646)	(1.615)	(0.997)	(1.387)	(1.230)
BPC			0.0282	0.1159	-0.0198	-0.0147	-0.0985
			(0.037)	(0.081)	(0.078)	(0.117)	(0.121)
NIM			-0.7549	-1.8988**	-0.0690	-0.1776	-0.3611
			(0.474)	(0.952)	(0.642)	(0.805)	(0.793)
BZS			0.1243	0.1915	0.2907*	0.2976	0.2150
			(0.093)	(0.191)	(0.170)	(0.188)	(0.208)
LogGDPperCapita					-2.3350	2.7033	-0.5727
					(3.354)	(3.800)	(4.136)
RealInterestrate					-0.5548***	-0.6066***	-0.5917***
					(0.167)	(0.216)	(0.209)
RealEffectiveExchangeRate					0.3665*	0.4020*	0.4181**
					(0.193)	(0.218)	(0.207)
InflationGDPdeflatorannual					-0.1459	-0.2051	-0.2442
					(0.203)	(0.268)	(0.252)
FINFIU					-0.1567	0.9056	1.1201
					(4.021)	(5.213)	(5.303)
VoiceandAccountability					3.8118	5.2576	4.0202
					(6.456)	(8.691)	(12.022)
PoliticalStabilityandAbsence					6.3691*	9.7714**	14.1852**
					(3.854)	(4.916)	(6.297)
GovernmentEffectiveness					2.7348	-3.3095	6.2568
					(9.130)	(11.606)	(14.665)
RegulatoryQuality					15.3472	8.5429	12.3506
					(11.197)	(14.997)	(14.513)
RuleofLaw					-30.4681**	-31.9085*	-42.7462**
					(11.893)	(17.024)	(17.990)
ControlofCorruption					20.5234*	26.0755*	19.1584
					(10.942)	(15.747)	(16.071)
Polity2							-0.0368
							(1.257)
Durable							0.3121***
							(0.101)
SizeofShadowEconomy							-0.0633
							(0.387)
Constant	7.4863	14.4233	0.5291	15.5096	-20.6367	-40.5395	-20.2945
	(6.410)	(18.696)	(6.875)	(17.200)	(35.105)	(44.644)	(44.738)
Observations	1,515	396	1,109	396	556	396	396
Number of ID	121	85	113	85	93	85	85
R-squared (overall)	0.00573	0.0134	0.0143	0.0381	0.0688	0.0858	0.102

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Table 3.b. Capital flows: Year fixed effects

VARIABLES	(1) BankFlowMeasure	(2) BankFlowMeasure	(3) BankFlowMeasure	(4) BankFlowMeasure	(5) BankFlowMeasure	(6) BankFlowMeasure	(7) BankFlowMeasure
FATF	-5.2146*	4.6324	-0.2241	10.4847	13.9006*	24.0445***	26.0158***
	(2.839)	(5.303)	(2.821)	(6.520)	(8.279)	(8.391)	(7.791)
OAR	-3.7818	-8.0745	0.2443	-4.4520	7.1562	3.6235	5.8246
	(4.876)	(9.033)	(3.598)	(7.448)	(5.729)	(8.239)	(7.529)
ISA	3.1768	9.6406*	3.7805*	7.2151	2.7590	3.4291	2.8250
	(2.128)	(5.051)	(2.050)	(4.657)	(3.287)	(4.043)	(3.631)
SLI	0.2646	0.2629	-0.3714	-0.2804	-0.6828	-1.5817	-1.0171
	(0.612)	(1.837)	(0.680)	(1.618)	(1.151)	(1.392)	(1.217)
BPC			0.0341	0.1146	-0.0354	-0.0419	-0.1253
			(0.035)	(0.078)	(0.077)	(0.113)	(0.117)
NIM			-0.8826*	-2.4725**	-0.1972	0.0744	-0.1438
			(0.517)	(1.060)	(0.791)	(1.146)	(1.129)
BZS			0.1167	0.1684	0.2305	0.2177	0.1366
			(0.091)	(0.210)	(0.164)	(0.200)	(0.219)
LogGDPperCapita					3.0276	7.2851*	3.7974
					(3.427)	(4.361)	(4.247)
RealInterestrate					-0.8961***	-1.0534***	-1.0151***
					(0.250)	(0.314)	(0.325)
RealEffectiveExchangeRate					0.2997	0.1939	0.2183
					(0.214)	(0.268)	(0.262)
InflationGDPdeflatorannual					-0.5203**	-0.7805*	-0.8016**
					(0.260)	(0.416)	(0.403)
FINFIU					5.1887	8.6972	8.8133*
					(4.330)	(5.686)	(5.303)
VoiceandAccountability					5.0865	5.0453	2.7320
					(6.333)	(8.328)	(13.475)
PoliticalStabilityandAbsence					4.9220	9.3907*	13.9066**
					(4.034)	(5.268)	(6.339)
GovernmentEffectiveness					-0.6031	-2.6515	6.9144
					(9.300)	(12.547)	(14.806)
RegulatoryQuality					6.5667	-2.3200	1.4325
					(11.010)	(16.094)	(15.607)
RuleofLaw					-25.6848**	-27.9471	-38.9196**
					(12.070)	(17.180)	(17.917)
ControlofCorruption					19.4581*	25.5137*	19.0729
					(10.719)	(14.979)	(15.122)
Polity2							0.1496
							(1.502)
Durable							0.3037***
							(0.082)
SizeofShadowEconomy							-0.1244
							(0.384)
Constant	-3.9061	12.1058	-4.6455	12.1644	-61.8431		-35.6012
	(6.391)	(18.851)	(6.991)	(17.834)	(39.772)		(50.486)
Observations	1,515	396	1,109	396	556	396	396
Number of ID	121	85	113	85	93	85	85
Country FE	NO	NO	NO	NO	NO	NO	NO
Year FE	YES	YES	YES	YES	YES	YES	YES
R-squared (overall)	0.149	0.0998	0.112	0.131	0.175	0.190	0.206

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Table 3.c. Capital flows: Year and Country fixed effects

VARIABLES	(1) BankFlowMeasure	(2) BankFlowMeasure	(3) BankFlowMeasure	(4) BankFlowMeasure
FATF	9.2667**	9.1494**	18.6461**	37.4726***
	(3.697)	(4.041)	(9.249)	(12.537)
OAR	4.3841	6.2674*	9.7371*	7.2191
	(6.636)	(3.668)	(5.787)	(8.608)
ISA	0.0760	3.0006	3.2043	11.5833*
	(2.551)	(2.930)	(5.599)	(6.612)
SLI	-0.7600	-0.6554	-1.6707	-3.5938**
	(0.714)	(0.871)	(1.235)	(1.504)
BPC		-0.4790*	-0.4328	-1.8172**
		(0.266)	(0.345)	(0.727)
NIM		0.0895	0.4856	-0.2752
		(1.060)	(0.912)	(1.430)
BZS		0.1577	-0.3247	-0.8322
		(0.300)	(0.517)	(0.755)
LogGDPperCapita			44.6581**	74.4295**
			(18.379)	(32.545)
RealInterestrate			-0.9874*	-0.2746
			(0.527)	(0.833)
RealEffectiveExchangeRate			0.0697	0.0807
			(0.342)	(0.502)
InflationGDPdeflatorannual			-0.7476**	-0.2973
			(0.363)	(0.583)
FINFIU			13.6012	11.4278
			(18.118)	(19.295)
VoiceandAccountability			2.5225	0.2718
			(17.453)	(24.094)
PoliticalStabilityandAbsence			16.0631	26.9338
			(20.694)	(29.186)
GovernmentEffectiveness			31.9474	28.9545
			(24.454)	(33.622)
RegulatoryQuality			-2.7553	0.3891
			(22.718)	(34.588)
RuleofLaw			-23.1351	4.8749
			(27.521)	(34.933)
ControlofCorruption			39.1591*	19.6466
			(20.329)	(29.555)
Polity2				-0.5771
				(4.647)
Durable				-1.8905
				(2.396)
SizeofShadowEconomy				-7.5073*
				(4.198)
Constant	-0.1269	26.7610	-402.6403***	-255.8468
	(8.119)	(21.346)	(141.124)	(284.460)
Observations	1,515	1,109	556	396
R-squared	0.162	0.134	0.207	0.290
Number of ID	121	113	93	85
Country FE	YES	YES	YES	YES
Year FE	YES	YES	YES	YES

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

List of the 85 countries considered in the last regression.

	Country	Years
1	Argentine	2004
2	Algeria	1999 and 2001-2006
3	Angola	2005
4	Australia	1999 and 2001-2007
5	Bahrain	1999 and 2001-2005
6	Bangladesh	2005
7	Belarus	2005
8	Belgium	2001-2007
9	Bhutan	2005
10	Bolivia	1999 and 2001-2006
11	Botswana	2005
12	Brazil	1999 and 2001-2006
13	Bulgaria	1999 and 2001-2007
14	Burundi	1999 and 2001-2007
15	Canada	1999 and 2001-2007
16	China	2003-2006
17	Chile	1999 and 2001-2006
18	Colombia	1999 and 2001-2006
19	Croatia	1999 and 2001-2007
20	Costa Rica	1999 and 2001-2006
21	Cyprus	1999 and 2001-2007
22	Czech Republic	1999 and 2001-2006
23	Denmark	1999 and 2001-2002
24	Egypt	2005
25	Dominican Republic	2003-2006
26	Estonia	2005
27	Ecuador	1999 and 2001-2002
28	Finland	2001-2004
29	France	2001-2004
30	Germany	2001-2002
31	Greece	1999 and 2002-2003
32	Guatemala	2004

33	Guyana	1999 and 2001-2006
34	Hungary	1999 and 2001-2007
35	Honduras	2004
36	India	2005
37	Indonesia	2005
38	Ireland	2001-2005
39	Israel	1999 and 2001-2007
40	Italy	2001-2007
41	Japan	1999 and 2001-2006
42	Jordan	2005
43	Kenya	2005
44	South Korea	2005
45	Kuwait	2005
46	Kyrgyzstan	2005
47	Latvia	2005
48	Lebanon	2005
49	Lesotho	1999 and 2001-2007
50	Macedonia	1999 and 2001-2006
51	Malaysia	1999 and 2001-2006
52	Mauritius	2005
53	Mexico	1999 and 2001-2006
54	Moldova	1999 and 2001-2006
55	Morocco	1999 and 2001-2005
56	Netherlands	2001-2007
57	New Zealand	1999 and 2001-2007
58	Nicaragua	1999 and 2001-2006
59	Nigeria	1999 and 2001-2006
60	Oman	2005
61	Pakistan	2004-2007
62	Panama	2004
63	Philippines	1999 and 2001-2007
64	Paraguay	1999 and 2001-2002
65	Poland	1999 and 2001-2006
66	Peru	2004
67	Romania	1999 and 2001-2007

68	Russia	1999 and 2001-2007
69	Singapore	1999 and 2001-2007
70	Slovakia	1999 and 2001-2007
71	Slovenia	2005
72	South Africa	1999 and 2001-2007
73	Spain	2001-2002
74	Sri Lanka	2005
75	Sweden	1999 and 2001-2005
76	Switzerland	1999 and 2001-2007
77	Suriname	2004
78	Tajikistan	2005
79	Thailand	2005
80	Trinidad and Tobago	1999 and 2001-2006
81	Ukraine	1999 and 2001-2002
82	United Kingdom	1999 and 2001-2007
83	United States	1999 and 2001-2007
84	Uruguay	1999 and 2001-2006
85	Venezuela	1999 and 2001-2006

Summary statistics for the 85 countries.

```
Variable         |      Mean   Std. Dev.       Min        Max |    Observations
-----------------+--------------------------------------------+----------------
BankFlow overall |  21.28501   64.45684    -141.02   554.2036 |     N =     396
         between |            30.43566  -40.34729    193.544 |     n =      85
         within  |            57.89944  -164.5387   484.1908 |     T = 4.65882
                 |                                            |
FATF     overall |  .0606061   .2389081          0          1 |     N =     396
         between |            .1688157          0          1 |     n =      85
         within  |            .1799555  -.6536797   .9356061 |     T = 4.65882
```

Table 4. Instrumental Variables

Variable	Description
Goldstein scale	l Level of International Cooperation of a country
Dummy cooperation	Based on the Goldstein Scale for the WEIS data
UN member	Membership in the UN Security Council
UN agenda	Whether the country has or has not been mentioned on the UN meeting agenda (http://research.un.org/en/docs/sc/quick)

Table 5. IV method: First stage

VARIABLES	(1) FATF	(2) FATF	(3) FATF	(4) FATF
OAR	-0.0850**	-0.1341**	-0.1374***	-0.1439***
	(0.035)	(0.053)	(0.051)	(0.054)
ISA	0.0137	0.0144	0.0115	0.0087
	(0.015)	(0.019)	(0.019)	(0.019)
SLI	-0.0064	-0.0139*	-0.0138*	-0.0136*
	(0.005)	(0.007)	(0.007)	(0.007)
BPC	-0.0005*	-0.0012**	-0.0012***	-0.0012**
	(0.000)	(0.000)	(0.000)	(0.000)
NIM	-0.0031	-0.0003	-0.0003	-0.0006
	(0.006)	(0.009)	(0.009)	(0.009)
BZS	-0.0003	-0.0012	-0.0015	-0.0015
	(0.002)	(0.002)	(0.002)	(0.002)
LogGDPperCapita	0.0099	-0.0308	-0.0322	-0.0294
	(0.018)	(0.028)	(0.028)	(0.028)
RealInterestrate	0.0002	0.0009	0.0011	0.0013
	(0.001)	(0.002)	(0.001)	(0.002)
RealEffectiveExchangeRate	-0.0016	0.0001	0.0001	-0.0001
	(0.001)	(0.002)	(0.002)	(0.002)
InflationGDPdeflatorannual	0.0002	0.0006	0.0006	0.0008
	(0.002)	(0.003)	(0.003)	(0.003)
FINFIU	-0.0200	-0.0531	-0.0556	-0.0536
	(0.029)	(0.046)	(0.045)	(0.045)
VoiceandAccountability	-0.0238	-0.1060	-0.0930	-0.0839
	(0.046)	(0.075)	(0.074)	(0.075)
PoliticalStabilityandAbsence	-0.0631**	-0.0664	-0.0701	-0.0777
	(0.032)	(0.051)	(0.050)	(0.050)
GovernmentEffectiveness	0.1021	0.3492***	0.3591***	0.3728***
	(0.063)	(0.104)	(0.104)	(0.102)
RegulatoryQuality	-0.0329	-0.0328	-0.0441	-0.0617
	(0.044)	(0.079)	(0.081)	(0.079)
RuleofLaw	0.1354*	0.1402	0.1397	0.1424
	(0.071)	(0.092)	(0.092)	(0.092)
ControlofCorruption	-0.1178***	-0.2225***	-0.2191***	-0.2228***
	(0.041)	(0.073)	(0.073)	(0.073)
Polity2	0.0024	0.0124	0.0109	0.0100
	(0.005)	(0.008)	(0.009)	(0.009)
Durable	-0.0002	-0.0005	-0.0006	-0.0007
	(0.000)	(0.000)	(0.001)	(0.001)
SizeofShadowEconomy	0.0060***	0.0080***	0.0079***	0.0077**
	(0.002)	(0.003)	(0.003)	(0.003)
UNSCmem	0.0348	-0.0134	-0.0339	-0.0298
	(0.033)	(0.054)	(0.056)	(0.056)
UNSCagenda	0.0201	0.0845	0.0864	0.0839
	(0.048)	(0.066)	(0.067)	(0.068)
GoldsteinCoop		0.2374**		-0.2148
		(0.114)		(0.237)
DummyCoop			1.5799***	2.4762*
			(0.591)	(1.260)
Constant	0.0077	0.2086	0.2485	0.2607
	(0.204)	(0.250)	(0.255)	(0.262)
Observations	396	253	253	253
R-squared	0.194	0.279	0.291	0.294
Country FE	NO	NO	NO	NO
Year FE	YES	YES	YES	YES

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Table 5.a. IV method: First stage using logit

VARIABLES	(1) FATF	(2) FATF	(3) FATF	(4) FATF	(5) FATF	(6) FATF	(7) FATF
OAR	-0.4296	-0.5847	-2.3098*	-9.3150**	-7.6865*	-20.7511**	-16.8345**
	(0.629)	(0.722)	(1.306)	(3.711)	(4.647)	(7.929)	(8.017)
ISA	0.0333	-0.2795	0.3598	1.6627	1.5705	-2.5963	-2.1330
	(0.314)	(0.373)	(0.585)	(2.215)	(2.629)	(4.309)	(4.271)
SLI	-0.0364	-0.0462	-0.0889	-0.7602	-1.4980*	-4.3865***	-3.9128***
	(0.090)	(0.107)	(0.179)	(0.574)	(0.827)	(1.481)	(1.511)
BPC		-0.0103	-0.0211	-0.2419***	-0.2675**	-0.4401***	-0.4173
		(0.010)	(0.019)	(0.079)	(0.114)	(0.150)	(0.267)
NIM		-0.0678	-0.1656	-0.6806	-2.4868**	-4.3162**	-4.1814**
		(0.119)	(0.215)	(0.658)	(1.266)	(1.927)	(2.043)
BZS		0.0321	-0.0117	-0.0025	0.2296	-0.4968	-0.4554
		(0.032)	(0.055)	(0.182)	(0.168)	(0.303)	(0.325)
LogGDPperCapita			-0.8445	-4.8165*	-1.7763	-10.1295*	-8.1236
			(0.605)	(2.521)	(3.035)	(6.154)	(6.012)
RealInterestrate			0.0211	0.0906	0.0421	0.0907	0.1191
			(0.042)	(0.115)	(0.123)	(0.179)	(0.189)
RealEffectiveExchangeRate			-0.0253	-0.2298	-0.1903	-0.1405	-0.1185
			(0.036)	(0.165)	(0.212)	(0.267)	(0.283)
InflationGDPdeflatorannual			-0.0900	-0.3028	0.2570	0.2852	0.3726
			(0.065)	(0.276)	(0.285)	(0.373)	(0.403)
FINFIU			0.1235	7.1632**	3.0050	-4.4004	-3.7035
			(1.067)	(3.217)	(3.777)	(6.530)	(6.267)
VoiceandAccountability				2.8027	-0.8232	2.6499	1.2071
				(4.373)	(7.551)	(11.908)	(12.730)
PoliticalStabilityandAbsence				-6.8166**	12.2991***	27.4544***	24.2579***
				(2.743)	(4.581)	(7.832)	(9.286)
GovernmentEffectiveness				-1.5941	5.2764	45.5046***	45.7461**
				(6.907)	(9.639)	(15.494)	(19.495)
RegulatoryQuality				3.4602	6.3599	0.5178	-3.7872
				(9.106)	(8.673)	(11.532)	(11.579)
RuleofLaw				11.9293**	25.8252***	27.6510**	24.8665
				(5.805)	(9.730)	(13.814)	(17.360)
ControlofCorruption				-4.2038	-9.0515	-24.2627**	-23.5632*
				(5.123)	(6.188)	(9.783)	(12.285)
Polity2					0.1721	5.4080***	5.1280**
					(0.843)	(1.861)	(1.998)
Durable					-0.3769**	-0.6956***	-0.6021**
					(0.154)	(0.242)	(0.273)
SizeofShadowEconomy					0.9194***	1.4066***	1.2612**
					(0.305)	(0.460)	(0.575)
UNSCmem	-0.5278	-0.4865	-0.7457	-3.1639	-1.1206	-5.3373	-2.6745
	(0.765)	(0.794)	(1.324)	(4.199)	(5.506)	(7.827)	(6.822)
UNSCagenda	1.9772*	2.6940**	3.8406**	11.5120***	1.3842	-2.4147	-0.9969
	(1.083)	(1.217)	(1.855)	(4.404)	(7.306)	(9.683)	(9.712)
GoldsteinCoop						21.6952	
						(21.939)	
DummyCoop							104.4118
							(108.761)
Constant	-5.6905**	-4.9836**	5.2082	47.4766*	-4.6239	71.8872	52.2286

	*	*					
	(1.373)	(1.793)	(6.833)	(26.683)	(28.355)	(51.662)	(52.379)
Observations	947	842	435	386	283	208	208
Number of ID	120	112	92	92	85	61	61
Country FE	NO	NO	NO	NO	NO	NO	NO
Year FE	YES	YES	YES	YES	YES	YES	YES

Standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Table 6. IV method: Second stage

VARIABLES	(1) BankFlowMeasure	(2) BankFlowMeasure	(3) BankFlowMeasure	(4) BankFlowMeasure
FATF	494.8670	75.1416	24.1770	8.2433
	(486.838)	(158.777)	(125.745)	(118.820)
OAR	46.7429	17.2951	10.0612	7.7996
	(41.926)	(25.399)	(22.991)	(22.870)
ISA	-3.0134	3.6969	4.1588	4.3033
	(10.631)	(6.156)	(5.960)	(5.916)
SLI	1.8822	0.6692	-0.0443	-0.2674
	(4.033)	(2.865)	(2.532)	(2.481)
BPC	0.0843	-0.1808	-0.2142	-0.2246
	(0.237)	(0.190)	(0.182)	(0.181)
NIM	1.4615	0.3290	0.2417	0.2144
	(3.575)	(1.528)	(1.463)	(1.468)
BZS	0.2968	0.2371	0.2128	0.2052
	(0.754)	(0.352)	(0.335)	(0.336)
LogGDPperCapita	-0.9039	9.7362	8.6602	8.3238
	(9.759)	(7.845)	(7.587)	(7.646)
RealInterestrate	-1.1491*	-1.4456***	-1.3947***	-1.3788***
	(0.674)	(0.482)	(0.459)	(0.456)
RealEffectiveExchangeRate	0.9931	0.0205	0.0137	0.0115
	(0.943)	(0.385)	(0.384)	(0.387)
InflationGDPdeflatorannual	-0.9267	-1.6092**	-1.5409**	-1.5195**
	(0.811)	(0.671)	(0.647)	(0.647)
FINFIU	17.3363	19.2028	16.7884	16.0335
	(16.055)	(13.964)	(13.316)	(13.253)
VoiceandAccountability	16.4492	-0.1410	-5.6599	-7.3853
	(26.087)	(24.295)	(21.301)	(20.734)
PoliticalStabilityandAbsence	45.3212	25.6919	20.6713	19.1016
	(35.733)	(17.887)	(15.009)	(14.569)
GovernmentEffectiveness	-41.1064	-10.6043	8.0141	13.8350
	(54.488)	(63.488)	(50.145)	(46.704)
RegulatoryQuality	13.2151	-14.7682	-17.6242	-18.5171
	(24.941)	(26.069)	(25.122)	(24.842)
RuleofLaw	-99.0667	-53.2921	-45.7472	-43.3884
	(72.467)	(34.696)	(30.231)	(29.204)
ControlofCorruption	74.2989	48.1746	36.4873	32.8334
	(61.438)	(43.740)	(35.478)	(33.280)
Polity2	-1.0952	0.5839	1.2739	1.4896
	(3.087)	(2.814)	(2.437)	(2.372)
Durable	0.3687	0.5065*	0.4874*	0.4814*
	(0.239)	(0.267)	(0.269)	(0.272)
SizeofShadowEconomy	-2.8188	-0.3197	0.0413	0.1542
	(2.883)	(1.276)	(1.059)	(1.010)
Constant	-48.3016	-91.1564	-83.0718	-80.5442
	(109.250)	(76.390)	(76.092)	(76.798)
Observations	396	253	253	253
R-squared	-2.236	0.265	0.279	0.273

Country FE	NO	NO	NO	NO
Year FE	YES	YES	YES	YES
R-squared (adjusted)	-2.483	0.184	0.200	0.193
Hansen J statistc	0.111	4.229	4.503	4.613
J statistic P-val	0.739	0.121	0.105	0.202
Underidentification P-val	0.509	0.139	0.0650	0.119

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Endogeneity test: FATF variable

	(1)	(2)	(3)	(4)
Endogeneity test	2.254	0.186	0.822	1.261
Chi-sq(1) P-val	0.1332	0.6659	0.3646	0.2614

Table 6.a. IV method: Second stage

VARIABLES	(1) BankFlowMeasure	(2) BankFlowMeasure	(3) BankFlowMeasure	(4) BankFlowMeasure
FATF	772.5685	58.7390	72.4696	61.3213
	(1,430.850)	(279.794)	(192.867)	(205.568)
OAR	70.9789	14.9669	16.9158	15.3334
	(118.462)	(43.980)	(32.438)	(34.266)
ISA	-6.4715	3.8456	3.7211	3.8222
	(21.503)	(6.027)	(5.781)	(5.797)
SLI	3.5995	0.4396	0.6318	0.4757
	(9.550)	(4.212)	(3.211)	(3.344)
BPC	0.2084	-0.1916	-0.1826	-0.1899
	(0.686)	(0.243)	(0.204)	(0.209)
NIM	2.4123	0.3009	0.3244	0.3053
	(7.560)	(1.601)	(1.568)	(1.556)
BZS	0.3917	0.2293	0.2358	0.2305
	(1.199)	(0.367)	(0.361)	(0.357)
LogGDPperCapita	-3.6885	9.3899	9.6798	9.4444
	(20.498)	(8.936)	(8.198)	(8.244)
RealInterestrate	-1.2284	-1.4292***	-1.4429***	-1.4318***
	(1.119)	(0.521)	(0.490)	(0.491)
RealEffectiveExchangeRate	1.4519	0.0183	0.0202	0.0187
	(2.322)	(0.390)	(0.387)	(0.386)
InflationGDPdeflatorannual	-1.0008	-1.5872**	-1.6056**	-1.5907**
	(1.292)	(0.724)	(0.685)	(0.685)
FINFIU	22.3846	18.4258	19.0763	18.5481
	(33.034)	(19.945)	(16.552)	(17.038)
VoiceandAccountability	24.5739	-1.9172	-0.4304	-1.6376
	(59.121)	(33.186)	(25.966)	(26.778)
PoliticalStabilityandAbsence	63.9281	24.0761	25.4287	24.3305
	(102.224)	(29.214)	(21.513)	(22.541)
GovernmentEffectiveness	-69.5493	-4.6121	-9.6282	-5.5555
	(147.615)	(103.009)	(70.750)	(75.345)
RegulatoryQuality	20.1940	-15.6874	-14.9180	-15.5427
	(48.672)	(30.915)	(26.864)	(27.492)
RuleofLaw	-134.6920	-50.8638	-52.8966	-51.2461
	(198.491)	(50.101)	(39.079)	(40.474)
ControlofCorruption	107.0094	44.4131	47.5618	45.0053
	(170.351)	(65.684)	(46.832)	(49.334)
Polity2	-1.8325	0.8059	0.6200	0.7710
	(6.571)	(4.103)	(3.129)	(3.248)
Durable	0.4072	0.5004*	0.5055*	0.5013*
	(0.383)	(0.298)	(0.285)	(0.287)
SizeofShadowEconomy	-4.4147	-0.2035	-0.3008	-0.2218
	(7.813)	(2.124)	(1.510)	(1.601)
Constant	-55.8241	-88.5544	-90.7325	-88.9641
	(159.057)	(89.966)	(83.418)	(84.235)
Observations	396	253	253	253
R-squared	-5.985	0.275	0.267	0.274
Country FE	NO	NO	NO	NO
Year FE	YES	YES	YES	YES
R-squared (adjusted)	-6.518	0.195	0.186	0.194
Hansen J statistic	0	0	0	0
J statistic P-val	.	.	.	.
Underidentification P-val	0.575	0.0576	0.0172	0.0221

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Endogeneity test: FATF variable

	(1)	(2)	(3)	(4)
Endogeneity test	2.441	0.007	0.037	0.016
Chi-sq(1) P-val	0.1182	0.9336	0.8477	0.8996

Table 7. External and Internal Bank Outflows

VARIABLES	(1) ExternalOutflow	(2) ExternalOutflows	(3) ExternalOutflows	(4) ExternalOutflows
FATF	0.0200	0.0247	0.0920	0.0626
	(0.059)	(0.073)	(0.095)	(0.097)
OAR	0.0595	0.0467	0.1047	0.0964
	(0.041)	(0.047)	(0.074)	(0.069)
ISA	0.0406*	0.0288	-0.0212	-0.0301
	(0.021)	(0.022)	(0.034)	(0.040)
SLI	0.0106	0.0030	-0.0008	-0.0113
	(0.008)	(0.009)	(0.014)	(0.017)
BPC		-0.0009	0.0001	-0.0006
		(0.001)	(0.001)	(0.001)
NIM		-0.0075	0.0122	0.0139
		(0.009)	(0.013)	(0.019)
BZS		0.0049	-0.0006	-0.0028
		(0.003)	(0.005)	(0.005)
LogGDPperCapita			0.0128	0.2317*
			(0.067)	(0.128)
RealInterestrate			-0.0107**	-0.0070
			(0.005)	(0.004)
RealEffectiveExchangeRate			0.0055**	0.0034
			(0.002)	(0.003)
InflationGDPdeflatorannual			-0.0053	-0.0045
			(0.005)	(0.004)
FINFIU			0.0858	0.0699
			(0.073)	(0.071)
VoiceandAccountability			-0.1556	-0.2352
			(0.155)	(0.150)
PoliticalStabilityandAbsence			0.0087	-0.0556
			(0.092)	(0.112)
GovernmentEffectiveness			-0.0308	-0.2281
			(0.252)	(0.309)
RegulatoryQuality			0.1184	0.2158
			(0.200)	(0.218)
RuleofLaw			0.1247	0.4937*
			(0.234)	(0.291)
ControlofCorruption			0.0490	0.0280
			(0.151)	(0.163)
Polity2				-0.0392
				(0.024)
Durable				0.0134
				(0.012)
SizeofShadowEconomy				0.0211
				(0.021)
Constant	-0.0832	0.0076	-0.6231	-3.0614**
	(0.072)	(0.117)	(0.611)	(1.333)
Observations	1,412	1,080	546	388
R-squared	0.006	0.007	0.055	0.110
Number of ID	121	113	90	82
Country FE	YES	YES	YES	YES
Year FE	NO	NO	NO	NO

Robust standard errors in parentheses
*** p<0.01,** p<0.05,* p<0.1

Table 8. Bank Inflow, estimation Results

VARIABLES	(1) ExternalInflows	(2) ExternalInflows	(3) ExternalInflows	(4) ExternalInflows
FATF	0.0078	-0.0042	0.1319	0.1364
	(0.056)	(0.058)	(0.085)	(0.101)
OAR	0.0057	0.0468	0.0965	0.1276**
	(0.036)	(0.045)	(0.093)	(0.053)
ISA	-0.0234	-0.0310	-0.0889**	-0.0578*
	(0.016)	(0.019)	(0.034)	(0.034)
SLI	0.0079	0.0206**	0.0246	-0.0172
	(0.007)	(0.010)	(0.019)	(0.012)
BPC		-0.0015*	-0.0002	-0.0033
		(0.001)	(0.001)	(0.003)
NIM		0.0062	0.0357	0.0353
		(0.013)	(0.028)	(0.035)
BZS		0.0040	-0.0022	-0.0031
		(0.004)	(0.006)	(0.007)
LogGDPperCapita			-0.0586	-0.0285
			(0.078)	(0.158)
RealInterestrate			-0.0071**	-0.0017
			(0.003)	(0.004)
RealEffectiveExchangeRate			0.0028	0.0030
			(0.003)	(0.004)
InflationGDPdeflatorannual			-0.0021	0.0048
			(0.005)	(0.007)
FINFIU			0.1886*	0.1807
			(0.097)	(0.122)
VoiceandAccountability			-0.0810	-0.0195
			(0.203)	(0.198)
PoliticalStabilityandAbsence			-0.0954	-0.0956
			(0.112)	(0.089)
GovernmentEffectiveness			0.0818	-0.0649
			(0.208)	(0.190)
RegulatoryQuality			-0.0394	0.1821
			(0.250)	(0.301)
RuleofLaw			0.2608	0.4859**
			(0.171)	(0.242)
ControlofCorruption			-0.1468	-0.2027
			(0.185)	(0.250)
Polity2				-0.0322
				(0.050)
Durable				0.0149
				(0.015)
SizeofShadowEconomy				-0.0432*
				(0.022)
Constant	0.0821	0.0141	0.1729	0.9630
	(0.061)	(0.125)	(0.700)	(1.386)
Observations	1,448	1,108	555	396
R-squared	0.003	0.019	0.065	0.120
Number of ID	121	113	93	85
Country FE	YES	YES	YES	YES
Year FE	NO	NO	NO	NO

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Table 9.a. Capital flows: considering lag1, Random effects

VARIABLES	(1) BankFlowMeasure	(2) BankFlowMeasure	(3) BankFlowMeasure	(4) BankFlowMeasure	(5) BankFlowMeasure	(6) BankFlowMeasure	(7) BankFlowMeasure
L.BankFlowMeasure	-0.0350	0.2747**	0.0214	0.2257*	0.0739	0.1545	0.1702
	(0.022)	(0.135)	(0.032)	(0.130)	(0.083)	(0.139)	(0.135)
FATF	-16.5472***	-7.2641	-6.1297**	-2.4173	4.8034	9.8947	11.2474**
	(3.537)	(4.746)	(3.006)	(5.356)	(7.524)	(6.333)	(5.675)
OAR	1.6943	-8.0466	5.1010	-4.8758	9.8754*	0.7885	3.0616
	(5.218)	(8.230)	(3.786)	(7.001)	(5.561)	(8.112)	(7.391)
ISA	2.8055	7.3040	3.7406*	5.2229	1.7348	1.4169	0.8682
	(2.269)	(4.561)	(1.997)	(4.277)	(3.098)	(4.052)	(3.547)
SLI	1.6625**	0.2905	0.5684	-0.0966	0.5326	-1.3630	-0.8253
	(0.648)	(1.785)	(0.644)	(1.596)	(0.987)	(1.376)	(1.191)
BPC			0.0255	0.1145	-0.0206	-0.0156	-0.1020
			(0.037)	(0.078)	(0.078)	(0.117)	(0.121)
NIM			-0.7629	-1.6309*	-0.0057	-0.0212	-0.2195
			(0.469)	(0.931)	(0.616)	(0.748)	(0.741)
BZS			0.1243	0.1687	0.2871*	0.2780	0.1964
			(0.092)	(0.182)	(0.172)	(0.186)	(0.199)
LogGDPperCapita					-2.4244	2.2320	-1.1223
					(3.276)	(3.707)	(4.033)
RealInterestrate					-0.5479***	-0.5850***	-0.5665***
					(0.162)	(0.205)	(0.193)
RealEffectiveExchangeRate					0.3475*	0.3646	0.3793*
					(0.197)	(0.226)	(0.216)
InflationGDPdeflatorannual					-0.1411	-0.2031	-0.2382
					(0.197)	(0.260)	(0.245)
FINFIU					-0.1917	0.4306	0.5140
					(3.949)	(5.125)	(5.024)
VoiceandAccountability					3.5527	4.6867	2.4178
					(6.308)	(8.399)	(11.690)
PoliticalStabilityandAbsence					6.0663	9.2032*	13.7365**
					(3.852)	(4.848)	(6.237)
GovernmentEffectiveness					3.2120	-1.4693	8.4574
					(8.736)	(10.526)	(13.546)
RegulatoryQuality					14.0582	5.1430	8.6549
					(10.872)	(14.183)	(13.293)
RuleofLaw					-29.6360**	-30.7141*	-41.3823**
					(11.880)	(16.752)	(17.802)
ControlofCorruption					20.7442*	27.1307*	20.3168
					(10.833)	(15.596)	(15.906)
Polity2							0.0897
							(1.182)
Durable							0.3182***
							(0.100)
SizeofShadowEconomy							-0.0619
							(0.356)
Constant	4.8020	10.7965	0.6345	11.5254	-18.5250	-34.7610	-14.6323
	(6.951)	(17.551)	(6.670)	(15.782)	(35.127)	(45.213)	(44.212)
Observations	1,410	396	1,107	396	556	396	396
Number of ID	121	85	113	85	93	85	85
Country FE	NO	NO	NO	NO	NO	NO	NO

Year FE	NO	NO	NO	NO	NO	NO	NO
R-squared (overall)	0.00789	0.0251	0.0148	0.0458	0.0700	0.0889	0.106

Robust standard errors in parentheses, *** p<0.01, ** p<0.05, * p<0.1

Table 9.b. Capital flows: considering lag1, Year fixed effects

VARIABLES	(1) BankFlowMeasure	(2) BankFlowMeasure	(3) BankFlowMeasure	(4) BankFlowMeasure	(5) BankFlowMeasure	(6) BankFlowMeasure	(7) BankFlowMeasure
L.BankFlowMeasure	-0.0162	0.3062**	0.0233	0.2513*	0.0649	0.1796	0.1942
	(0.021)	(0.135)	(0.029)	(0.132)	(0.086)	(0.135)	(0.133)
FATF	-5.2561*	4.4938	-0.0908	10.1209*	13.6431*	23.2463***	25.0667***
	(2.895)	(5.070)	(2.762)	(6.030)	(8.135)	(8.105)	(7.491)
OAR	-1.5322	-7.1152	0.3426	-3.9334	7.0504	3.8049	6.0424
	(5.128)	(8.651)	(3.525)	(7.217)	(5.696)	(8.089)	(7.373)
ISA	3.5054	9.4930*	3.7276*	7.2430	2.8045	3.4074	2.8814
	(2.312)	(4.931)	(2.073)	(4.498)	(3.225)	(3.960)	(3.472)
SLI	0.3023	0.2150	-0.4066	-0.2728	-0.6965	-1.5446	-0.9924
	(0.639)	(1.807)	(0.672)	(1.599)	(1.142)	(1.367)	(1.169)
BPC			0.0320	0.1138	-0.0353	-0.0419	-0.1281
			(0.035)	(0.075)	(0.076)	(0.112)	(0.115)
NIM			-0.8846*	-2.1903**	-0.1543	0.2279	-0.0302
			(0.511)	(1.034)	(0.774)	(1.110)	(1.102)
BZS			0.1140	0.1448	0.2295	0.1964	0.1199
			(0.089)	(0.203)	(0.164)	(0.199)	(0.210)
LogGDPperCapita					3.0187	7.0322*	3.5469
					(3.370)	(4.245)	(4.100)
RealInterestrate					-0.8983***	-1.0415***	-1.0008***
					(0.252)	(0.315)	(0.323)
RealEffectiveExchangeRate					0.2846	0.1551	0.1810
					(0.217)	(0.270)	(0.265)
InflationGDPdeflatorannual					-0.5164**	-0.7820*	-0.7961**
					(0.260)	(0.408)	(0.396)
FINFIU					5.2573	8.3423	8.3120*
					(4.256)	(5.518)	(4.952)
VoiceandAccountability					4.8281	4.2808	0.2788
					(6.169)	(7.881)	(13.183)
PoliticalStabilityandAbsence					4.6633	8.6962*	13.3831**
					(4.018)	(5.099)	(6.183)
GovernmentEffectiveness					-0.4258	-1.0449	8.7351
					(9.016)	(11.651)	(13.922)
RegulatoryQuality					5.5181	-6.3637	-2.9214
					(10.757)	(15.393)	(14.579)
RuleofLaw					-25.0875**	-26.3151	-36.8940**
					(12.010)	(16.758)	(17.606)
ControlofCorruption					19.8064*	26.6613*	20.4153
					(10.644)	(14.737)	(14.848)
Polity2							0.3641
							(1.449)
Durable							0.3089***
							(0.081)
SizeofShadowEconomy							-0.1227
							(0.353)
Constant	-6.1417		-4.3198	4.2572	-59.9255		-35.2721
	(6.775)		(6.909)	(15.595)	(40.038)		(48.467)
Observations	1,410	396	1,107	396	556	396	396
Number of ID	121	85	113	85	93	85	85

Country FE	NO	NO	NO	NO	NO	NO	NO
Year FE	NO	NO	NO	NO	NO	NO	NO
R-squared (overall)	0.154	0.114	0.113	0.140	0.176	0.194	0.211

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Table 9.c. Capital flows: Arellano-Bond estimators

VARIABLES	(1) BankFlowMeasure	(2) BankFlowMeasure	(3) BankFlowMeasure	(4) BankFlowMeasure
L.BankFlowMeasure	-0.0732***	-0.0096	-0.1090**	-0.0630
	(0.022)	(0.038)	(0.054)	(0.092)
FATF	2.8305	4.6796	-2.6638	0.9113
	(6.035)	(4.464)	(5.347)	(4.872)
OAR	7.8053	15.4517***	20.4588***	19.7741***
	(5.481)	(5.370)	(7.327)	(6.633)
ISA	-1.1622	-1.7618	-6.0147	0.0899
	(3.877)	(4.003)	(5.255)	(4.233)
SLI	-1.0211	-0.4191	0.8057	-1.3648
	(1.028)	(1.095)	(1.463)	(1.824)
BPC		-0.3600	-0.0739	-0.0270
		(0.370)	(0.087)	(0.181)
NIM		1.0716	-0.0330	-1.4389
		(0.905)	(0.909)	(1.056)
BZS		0.5139	0.7689*	0.7139
		(0.374)	(0.412)	(0.464)
LogGDPperCapita			18.3849	-15.0106
			(19.128)	(15.076)
RealInterestrate			-0.6050**	-0.4705
			(0.240)	(0.326)
RealEffectiveExchangeRate			0.3916	0.9100***
			(0.306)	(0.243)
InflationGDPdeflatorannual			-0.3981**	-0.4880**
			(0.183)	(0.220)
FINFIU			-2.9049	-6.5432
			(5.513)	(5.662)
VoiceandAccountability			3.6154	-3.9253
			(18.659)	(19.440)
PoliticalStabilityandAbsence			-14.3267	-13.4908
			(13.221)	(14.174)
GovernmentEffectiveness			10.8511	31.4350
			(22.700)	(23.284)
RegulatoryQuality			1.6588	7.6260
			(12.287)	(14.893)
RuleofLaw			-11.5681	19.5476
			(32.823)	(36.488)
ControlofCorruption			12.2141	-3.8521
			(11.716)	(13.732)
Polity2				-2.2282
				(4.628)
Durable				-0.5899
				(3.050)
SizeofShadowEconomy				-1.5472
				(4.421)
Constant	96.2568***	5.7964	-213.0330	128.3888
	(20.297)	(27.929)	(156.496)	(277.533)
Observations	1,276	986	411	265
Number of ID	121	111	62	57
Year FE	YES	YES	YES	YES

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Table 9.1.a. Capital flows: considering lag1 of bank inflows, Random effects

VARIABLES	(1) BankFlowMeasure	(2) BankFlowMeasure	(3) BankFlowMeasure	(4) BankFlowMeasure
L.ExternalInflows	3.9573	4.0214	4.8758	8.4790*
	(3.501)	(3.714)	(4.994)	(5.131)
FATF	-14.9434***	-5.8340**	4.9719	11.4957**
	(3.246)	(2.916)	(7.309)	(5.707)
OAR	1.6663	4.9911	10.3559*	3.4135
	(5.247)	(3.848)	(5.553)	(7.530)
ISA	2.5437	3.6586*	1.6006	0.6643
	(2.130)	(1.980)	(3.121)	(3.590)
SLI	1.3834**	0.5304	0.4724	-0.8058
	(0.643)	(0.648)	(1.004)	(1.239)
BPC		0.0301	-0.0209	-0.0961
		(0.037)	(0.078)	(0.120)
NIM		-0.7152	-0.0404	-0.3578
		(0.475)	(0.631)	(0.791)
BZS		0.1300	0.3002*	0.2313
		(0.092)	(0.169)	(0.207)
LogGDPperCapita			-2.4264	-1.4399
			(3.336)	(4.089)
RealInterestrate			-0.5442***	-0.5836***
			(0.163)	(0.205)
RealEffectiveExchangeRate			0.3541*	0.4234**
			(0.192)	(0.213)
InflationGDPdeflatorannual			-0.1192	-0.1851
			(0.192)	(0.238)
FINFIU			-0.0346	1.1862
			(3.980)	(5.218)
VoiceandAccountability			3.7321	4.7625
			(6.548)	(11.754)
PoliticalStabilityandAbsence			6.1160	14.3129**
			(3.854)	(6.267)
GovernmentEffectiveness			3.4159	6.1344
			(9.120)	(14.665)
RegulatoryQuality			15.7650	12.5748
			(11.274)	(14.360)
RuleofLaw			-31.4401***	-43.2355**
			(12.150)	(17.893)
ControlofCorruption			20.9741*	19.9982
			(11.008)	(16.275)
Polity2				-0.0887
				(1.233)
Durable				0.3142***
				(0.101)
SizeofShadowEconomy				-0.0692
				(0.376)
Constant	4.9942	0.3188	-19.3765	-15.3910
	(6.759)	(6.864)	(35.280)	(45.544)
Observations	1,343	1,105	555	396
Number of ID	121	113	93	85
Country FE	NO	NO	NO	NO
Year FE	NO	NO	NO	NO
R-squared (overall)	0.00595	0.0154	0.0710	0.105

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Table 9.1.b. Capital flows: considering lag1 of bank inflows, Year fixed effects

VARIABLES	(1) BankFlowMeasure	(2) BankFlowMeasure	(3) BankFlowMeasure	(4) BankFlowMeasure
ExternalInflows_1	8.3984**	5.6238	7.7260	15.7410*
	(4.028)	(3.747)	(6.019)	(9.074)
FATF	-4.6252*	0.0686	13.7990*	25.4849***
	(2.629)	(2.699)	(7.752)	(7.249)
OAR	-2.3864	0.1913	6.9905	6.8544
	(5.293)	(3.575)	(5.607)	(7.537)
ISA	3.6627	3.7577*	2.7929	2.2861
	(2.228)	(2.065)	(3.291)	(3.581)
SLI	0.1543	-0.4003	-0.6526	-0.8872
	(0.662)	(0.678)	(1.150)	(1.208)
BPC		0.0366	-0.0359	-0.1202
		(0.035)	(0.076)	(0.113)
NIM		-0.8368	-0.1621	-0.2256
		(0.518)	(0.809)	(1.162)
BZS		0.1207	0.2468	0.1658
		(0.089)	(0.164)	(0.215)
LogGDPperCapita			3.0098	2.5124
			(3.457)	(4.265)
RealInterestrate			-0.8924***	-1.0120***
			(0.250)	(0.316)
RealEffectiveExchangeRate			0.2925	0.2304
			(0.210)	(0.265)
InflationGDPdeflatorannual			-0.4865*	-0.6983*
			(0.250)	(0.372)
FINFIU			5.3729	9.2912*
			(4.240)	(5.146)
VoiceandAccountability			5.0866	3.2249
			(6.251)	(12.998)
PoliticalStabilityandAbsence			4.6608	14.1223**
			(4.009)	(6.365)
GovernmentEffectiveness			-0.8974	6.1004
			(9.148)	(14.683)
RegulatoryQuality			6.4217	1.4252
			(11.042)	(15.396)
RuleofLaw			-24.8859**	-38.9209**
			(11.834)	(17.716)
ControlofCorruption			19.3720*	20.4567
			(10.562)	(15.350)
Polity2				0.1509
				(1.479)
Durable				0.3049***
				(0.081)
SizeofShadowEconomy				-0.1486
				(0.371)
Constant	-5.8353	-5.2845	-62.1114	-30.5972
	(6.818)	(6.993)	(40.629)	(51.720)
Observations	1,343	1,105	555	396
Number of ID	121	113	93	85
Country FE	NO	NO	NO	NO
Year FE	YES	YES	YES	YES
R-squared (overall)	0.159	0.114	0.179	0.214

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Table 9.1.c. Capital flows: considering lag1 of bank inflows, Year and Country fixed effects

	(1)	(2)	(3)	(4)
VARIABLES	BankFlowMeasure	BankFlowMeasure	BankFlowMeasure	BankFlowMeasure
ExternalInflows_1	7.8601**	4.7398	4.6140	7.1847
	(3.562)	(3.591)	(6.189)	(8.919)
FATF	8.3310**	9.3352**	18.5592**	37.8459***
	(3.694)	(4.027)	(9.250)	(12.544)
OAR	5.9041	6.3258*	9.3194*	6.6553
	(7.532)	(3.593)	(5.603)	(8.728)
ISA	-0.5090	3.3225	3.4479	11.4208*
	(2.833)	(2.920)	(5.642)	(6.637)
SLI	-0.9361	-0.7396	-1.6656	-3.4985**
	(0.795)	(0.872)	(1.213)	(1.503)
BPC		-0.4778*	-0.4314	-1.7988**
		(0.264)	(0.343)	(0.725)
NIM		0.1542	0.5371	-0.0955
		(1.037)	(0.908)	(1.499)
BZS		0.1752	-0.3149	-0.7989
		(0.300)	(0.519)	(0.753)
LogGDPperCapita			44.6710**	73.6039**
			(18.458)	(32.310)
RealInterestrate			-0.9974*	-0.3003
			(0.526)	(0.833)
RealEffectiveExchangeRate			0.0637	0.0691
			(0.337)	(0.496)
InflationGDPdeflatorannual			-0.7232**	-0.2262
			(0.364)	(0.591)
FINFIU			13.8606	12.2230
			(18.119)	(19.289)
VoiceandAccountability			1.8233	-0.7933
			(17.896)	(24.587)
PoliticalStabilityandAbsence			14.7766	26.6232
			(20.655)	(29.399)
GovernmentEffectiveness			32.8884	27.5464
			(23.894)	(34.650)
RegulatoryQuality			-2.9384	1.4519
			(22.844)	(34.280)
RuleofLaw			-21.8500	4.9601
			(27.622)	(34.859)
ControlofCorruption			38.4112*	21.4430
			(20.065)	(30.369)
Polity2				-0.4981
				(4.586)
Durable				-1.6724
				(2.341)
SizeofShadowEconomy				-7.6859*
				(4.199)
Constant	0.8098	25.6758	-403.3741***	-256.7684
	(9.219)	(21.218)	(142.009)	(283.452)
Observations	1,343	1,105	555	396
R-squared	0.170	0.135	0.208	0.292
Number of ID	121	113	93	85
Country FE	YES	YES	YES	YES
Year FE	YES	YES	YES	YES

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Appendix 3

List of variables

DEPENDENT VARIABLES	DEFINITION AND INFORMATION	SOURCE
External Assets (all_sector)	Asset positions vis-à-vis banks and non-banks located in a country other than the country of residence of the reporting banking office.	Bank for International Settlements (BIS) Locational banking statistics
External Liabilities (all_sector)	Liability positions vis-à-vis banks and non-banks located in a country other than the country of residence of the reporting banking office.	Bank for International Settlements (BIS) Locational banking statistics
Bank Flow Measure	Annual percentage change in Total Foreign Claims, computed as 100*[Ln(TFC(t)) - Ln(TFC(t-1))]. Total Foreign Claims are the sum of Local Claims of foreign affiliates in local currency and Total International Claims, i.e. the sum of Local Claims of foreign affiliates in foreign currency and Cross-Border Claims (activities that are granted or extended to non-residents).	Bank for International Settlements (BIS) Consolidated banking statistics

CONTROL VARIABLES	TOPIC	DEFINITION AND INFORMATION	SOURCE
FATF BlackList	Anti-Money Laundering and Combatting the Financing of Terrorism Compliance	FATF BlackList is a Dummy variable that takes the value of 1 if the Country was blacklisted or monitored during the study period, 0 otherwise.	FATF Annual and Overall Reviews of Non-Cooperative Countries or Terrirories; FATF web site.
Overall Activities Restrictions Index	Bank Regulation and Supervision	Overall Activities Restrictions Index is a Dummy variable that takes the value of 1 if the banks of the countries are subject to some restrictions in their operations, 0 if they are totally free in their investment choices.	Barth J.R., Caprio G. Jr., Levine R. "Banking Regulation Surveys"
Independence of Supervisory Authority Index	Bank Regulation and Supervision	The Independence of Supervisory Authority Index is an indicator that takes the values of 0, 1, 2 or 3 and it expresses the degree to which the supervisory authority is independent from the government and legally protected from the banking industry.	Barth J.R., Caprio G. Jr., Levine R. "Banking Regulation Surveys"
Supervisory Lightness Index	Bank Regulation and Supervision	The Supervisory Lightness Index is an index which takes values between 0 and 20 and it's an inverse indicator of the degree of lightness of the supervision system. It's built up starting from 20 question of the Banking Regulation Surveys of Barth et al. The Supervisory Lightness Index is the total sum of the 20 indicators.	Barth J.R., Caprio G. Jr., Levine R. "Banking Regulation Survey".
Growth of GDP (annual %)	Macroeconomic Characteristic	Annual percentage growth rate of GDP at market prices based on constant local currency. Aggregates are based on constant 2000 U.S. dollars. GDP is the sum of gross value added by all resident producers in the economy plus any product taxes and minus any subsidies not included in the value of the products.	World Bank national accounts data, and OECD National Accounts data files.
Log GDP per Capita	Macroeconomic Characteristic	The natural logarithm of the GDP per capita. GDP per capita is gross domestic product divided by midyear population. GDP is the sum of gross value added by all resident producers in the economy plus any product taxes and minus any subsidies not included in the value of the products.	World Bank national accounts data, and OECD National Accounts data files.
Bank Private Credit to GDP (%)	Financial Institutions Depth (size)	The financial resources provided to the private sector by domestic money banks as a share of GDP. Domestic money banks comprise commercial banks and other financial institutions that accept transferable deposits, such as demand	International Financial Statistics (IFS) - International Monetary Fund (IMF)

		deposits.	
Net Interest Marginal (%)	Financial Institutions Efficiency	Bank’s income that has been generated by non-interest related activities as a percentage of total income (net-interest income plus non-interest income). Non-interest related income includes net gains on trading and derivatives, net gains on other securities, net fees and commissions and other operating income.	Bankscope
Bank Z Score	Financial Institutions Stability	It captures the probability of default of a country's banking system, calculated as a weighted average of the z-scores of a country's individual banks (the weights are based on the individual banks' total assets). Z-score compares a bank’s buffers (capitalization and returns) with the volatility of those returns.	Bankscope
FINFIU	Banking and Financial Supervision	It's a dummy variable that takes the value of 1 if a Financial Intelligence Unit exist and it has a financial nature (i.e. is under an authority or an entity involved in finacial markets), 0 otherwise.	Information taken from the FIUs' web site
Real Interest rate (%)	Macroeconomic Characteristic	Real interest rate is the lending interest rate adjusted for inflation as measured by the GDP deflator.	World Bank
Real Effective Exchange Rate	Macroeconomic Characteristic	The effective exchange rate is a weighted average of a basket of foreign currencies, and it can be viewed as an overall measure of the country's external competitiveness. The real effective exchange rate is the nominal effective exchange rate divided by a price deflator (they use the CPI). Index 2005 =100.	IFS, IMF data.
Inflation, GDP deflator (annual %)	Macroeconomic Characteristic	Inflation as measured by the annual growth rate of the GDP implicit deflator shows the rate of price change in the economy as a whole. The GDP implicit deflator is the ratio of GDP in current local currency to GDP in constant local currency.	World Bank
Voice and Accountability	Public Governance	It reflects perceptions of the extent to which a country's citizens are able to participate in selecting their government, as well as freedom of expression, freedom of association, and a free media.	The Worldwide Governance Indicators, Kaufmann et al. 2008, World Bank.
Political Stability and Absence of Violence	Public Governance	It reflects perceptions of the likelihood that the government will be destabilized or overthrown by unconstitutional or violent means, including politically-motivated violence and terrorism.	The Worldwide Governance Indicators, Kaufmann et al. 2008, World Bank.
Government Effectiveness	Public Governance	It reflects perceptions of the quality of public services, the quality of the civil service and the degree of its independence from political pressures, the quality of policy formulation and implementation, and the credibility of the government's commitment to such policies.	The Worldwide Governance Indicators, Kaufmann et al. 2008, World Bank.
Regulatory Quality	Public Governance	It reflects perceptions of the ability of the government to formulate and implement sound policies and regulations that permit and promote private sector development.	The Worldwide Governance Indicators, Kaufmann et al. 2008, World Bank.
Rule of Law	Public Governance	It reflects perceptions of the extent to which agents have confidence in and abide by the rules of society, and in particular the quality of contract enforcement, property rights, the police, and the courts, as well as the likelihood of crime and violence.	The Worldwide Governance Indicators, Kaufmann et al. 2008, World Bank.
Control of Corruption	Public Governance	It reflects perceptions of the extent to which public power is exercised for private gain, including both petty and grand forms of corruption, as well as "capture" of the state by elites and private interests.	The Worldwide Governance Indicators, Kaufmann et al. 2008, World Bank.
Polity 2	Political Control	This variable is a modified version of the POLITY variable added in order to facilitate the use of the POLITY regime measure in time-series analyses. It modifies the combined annual POLITY score by applying a simple treatment to convert instances of “standardized authority scores” (i.e., -66, -77, and -88) to conventional polity scores (i.e., within the range, -10 to	Polity IV Project, University of Maryland

		+10). The Polity score is computed by subtracting the Autocracy score from the Democracy score.	
Durable	Political Control	This variable indicates the number of years since the most recent regime change.	Polity IV Project, University of Maryland
Size of Shadow Economy	Shadow Economy	It's a measure of the Size of the Shadow Economy.	Shadow Economies All over the World: New Estimates for 162 Countries from 1999 to 2007 (Revised Version), F. Schneider, A. Buehn and C. E. Montenegro

IV VARIABLES	**DEFINITION AND INFORMATION**	**SOURCE**
Goldstein scale	It captures the level of International Cooperation of a country	Goldstein (1992)
Dummy cooperation	It is based on the Goldstein Scale for the WEIS data. It shows 1 if a country cooperate, 0 if a country does not cooperate.	Goldstein (1992)
UN member	Dummy variable indicating the membership of a country in the UN Security Council in year t	UN website
UN agenda	Whether the country has or has not been mentioned on the UN meeting agenda in year t	UN website (http://research.un.org/en/docs/sc/quick)

Network Effects on Stock Market Participation

O. Balakina and A. Parakhonyak

Abstract

This paper introduces an equilibrium model of stock market participation with a social network to study how information diffusion affect the decision to enter the stock market. In the model agents of different types exist within a social network structure and share information through social interactions. We show that the equilibrium participation level depends on the number of informed agents and the intensity of social interactions.

We provide an algorithm for finding the equilibrium stock market participation for a particular network structure and use Danish registry data for the period 2010 -2013 to test the hypothesis that connectivity affect stock market participation. Consistent with the predictions of the model we find that connectivity has a positive impact on stock market participation. Finally, we find that the model with social networks outperform standard models of stock market participation.

1 Introduction

The reason why only one-half of U.S. households invest in the stock market and less than 20 percent hold stocks or mutual funds outside retirement plans has been discussed in the finance literature since 1980s without arriving at a convincing explanation. Given that risk premium is positive, theoretical models of optimal investment portfolios predict that a risk averse agent always invests a positive share of her wealth in stock (e.g. Mehra and Prescott, 1985). The proposed solution to this problem has been to include a fixed cost of stock market participation (e.g. Vissing-Jorgensen, 2002), which implies that some households do not participate as the expected benefit is lower than the cost. However, most of empirical studies estimate costs of participation that are too low to account for the puzzle (see Khorunzhina, 2013, and citations within). A potential explanation is that both the theoretical and empirical studies attempting to explain this "limited stock market participation puzzle" has so far ignored the social aspect of financial investments.[1] Contrary to the assumption of these models, financial decisions are often taken after consulting with friends, family or professionals who themselves have information to share.

In this paper, we propose a model that incorporates social interactions and information-sharing between agents into a basic portfolio model. There are two types of agents in the model that differ with respect to their fixed costs of participation. Educated agents have low costs of participation, and are not affected by interactions with their peers. Non-Educated agents have high fixed costs, which they can lower by learning from informed peers within their social networks. We show that limited stock market participation can arise in equilibrium because of less intense information sharing, even if the risk premium is positive.

The implication of the theoretical model is that empirical estimates of the fixed costs of stock market participation are biased downwards in the presence of social network effects. To see why, consider estimating the unobserved fixed costs for three agents: agents L1 and L2 have low income, and agent H has high income. Assume that all agents have the same fixed cost of participating, but that agent H has a sufficiently high income to warrant investing in the stock market. Further, assume that agent L1 and H are in the same social network and that they share information. Agent L2 is not included in the same network. As agent H participates and shares information within his social network, agent L1 also participates because her initial fixed cost has been reduced by his

[1] One exception is Kaustia and Knüpfer (2012).

social network. Without considering the role that the social network had in spreading information to agent L1, empirical estimates includes a low-income agent participating, which biases the estimated costs downwards.

In our model, risk-averse agents exist in a one period economy and can transfer their initial wealth from the beginning of the period, $t = 0$, to the end of the period, $t = 1$ by investing in two types of assets: a risky and a risk-free asset. There are two types of agents: educated and non-educated, which refers to their education about financial markets. All agents pay a positive fixed cost at $t = 0$ to access the risky asset, where the fixed cost depends on agents type. Educated agents pay a fixed participation costs that is smaller than costs of non-educated agents. We further assume that all agents in the economy belong to a social network and that connectivity depends on the parameters of the wealth distribution: agents with similar wealth have a higher probability of being connected in the network compared to agents with different wealth. Non-educated agents obtain information about stock market from their educated peers who invests in stocks, which thereby decreases their participation costs. The more participating peers (higher connectivity) in one's network, the lower are an agent's fixed participation costs. In equilibrium therefore, higher connectivity between agents leads to higher stock market participation, all else equal. The implication is that low level of connectivity explains low participation even among wealthy agents. In addition, in equilibrium a high level of connectivity boosts participation in the economy even if there is a low level of financial education. To our knowledge, this paper represents the first attempt to solve a theoretical model of stock market participation in the economy with the social network. We define a unique static equilibria for stock market participation in the economy with a specific type of the network. In addition to that, we also define the highest achievable level of stock market participation.

We proceed to test the hypothesis developed in the model using Danish administrative data on the full population of households. Consistent the predictions from the model, we find that areas with low "connectivity", measured as a low number of links between individuals or a low intensity of the network, have a lower stock market participation rate, consistent with the prediction of the model. Using micro-data on the full population we construct social networks based on workplaces, place of study and geographical location, conditional on the agents being in the same age bracket, and estimate whether connectivity affects stock market participation. Our results indicate that connectivity has a positive and significant impact on participation, even after controlling for financial and demographic characteristics of individuals.

Our paper contributes to several strands different literatures. Our main contribution is for the theoretical finance literature that recurrently discusses several explanations of the limited stock market participation puzzle. This puzzle dates back to Arrow (1965), who shows that investors at every risk-tolerance level optimally hold risky stocks because the equity premium is positive.[2] This puzzle has received considerable attention from researchers in economics and finance (see e.g. Mankiw and Zeldes, 1991; Haliassos and Bertaut, 1995; Heaton and Lucas, 2000; Brav, Constantinides, and Geczy, 2002). We contribute to this literature by showing that social networks can help explain low levels of participation. In our model, low connectivity between agents prevents information sharing and as a result limits stock market participation. Furthermore, we show that low stock market participation among wealthy agents can be explained by low connectivity in their social network. Previous explanation for this phenonomen includes lack of stock market awareness (Hong, Kubik, and Stein, 2004; Guiso and Jappelli, 2005; Brown et al., 2008), nonstandard preferences like ambiguity aversion (Dow and Costa Werlang, 1992; Ang, Bekaert, and Liu, 2005; Cao, Wang, and Zhang, 2005; Epstein and Schneider, 2007) and education deficits (Campbell, 2006; Calvet, Campbell, and Sodini, 2007; Christiansen, Joensen, and Rangvid, 2008; Van Rooij, Lusardi, and Alessie, 2011; Guiso, Sapienza, and Zingales, 2008).

Our results also provide new insights into why the estimated fixed costs of stock market participation are biased downwards. The empirical literature attempting to estimate these costs typically state that the empirical estimates are too low to explain limited participation. For example, Vissing-Jorgensen (2002) finds that moderate fixed participation costs explain the non-participation of many U.S. households. However, given the magnitude of these costs (about $250), fixed costs are unlikely to explain the full degree of non-participation. Indeed, Mankiw and Zeldes (1991) and Heaton and Lucas (2000) argue that such fixed costs do not explain the rate of non-participation among the wealthy. Our model provides some intuition for why the above studies may have underestimated fixed costs.

Our paper is also related to the more recent strand of literature that studies how social interactions between agents affects financial decisions. Kaustia and Knüpfer (2012) show that peer stock portfolio performance significantly affects stock market participation. Bursztyn et al. (2014) use field experiments to show two channels of social influence on financial decisions: *social learning*

[2] Mehra and Prescott (1985) and Fama and French (2002) show that stock market participation can have a substantial effect on lifetime consumption.

and *social utility* ("Keeping up with Joneses" type of preferences). They show that both channels have statistically and economically significant effects on investment decisions. Changwony, Campbell, and Tabner (2014) chose a more general formulation of social interactions to investigate the separate and joint influences of social engagement on stock market participation, and find that socially engaged individuals are more likely to participate. They find that a weak tie (measured by social group involvement) has a positive effect on stock market participation whereas a strong tie (measured by frequency of talking to neighbors) has no effect. Patacchini and Rainone (2014) study the importance of social interactions for the adoption of financial products, and find that not all social contacts are equally important. The authors also show that there are different types of networks (strong and weak in their terminology) and the effect of these two types of networks is different.

Finally, this paper contributes to the literature on the importance of introducing the concept of social networks in theory.[3]

The rest of the paper is organized as follows. In Section 2 we present the general setting of the theoretical model, provide details about the network structure and discuss the solution algorithm and the resulting equilibria. In Section 3 we provide results from model simulations. In Section 4 we discuss the Danish administrative data and the empirical testing of the model. Section 5 concludes.

2 Model

2.1 General setting

In this section we introduce a one period, closed economy model that describes the financial behavior of an agent with a social network. In the beginning of the period agents in the economy allocate their wealth between a risk-free and a risky asset, and in the end of the period they consume the proceeds from the portfolio in the form of a non-durable consumption good.

The economy is populated by risk-averse agents with the same type of CRRA preferences. The utility function is:

$$U(W) = \frac{W^{1-\gamma}}{1-\gamma},\ \gamma > 0$$

where W defines the level of wealth of an agent in the end of the period, and γ defines the level of

[3]See Zenou (2012) for a substantial literature review on the theory of networks and on applications.

relative risk aversion of the agent. Agents have initial wealth w_i distributed as $F(\cdot)$, $w_i \sim F(\cdot)$.

There are two investment opportunities in the economy. An agent can invest in a risk-free asset with gross return R^f. For simplification, we assume that $R^f = 1$, i.e. the net return on the risk-free asset is zero, $r^f = 0$. The agent can also allocate part of her initial wealth to a risky asset. The net return on the risky asset r is a random variable with binomial distribution, such that

$$r = \begin{cases} r_u & \text{,with probability } \pi \\ r_d & \text{,with probability } (1-\pi) \end{cases}$$

where $r_d < 0 < r_u$. The expected excess return on the stock market portfolio is positive, i.e.:

$$\pi r_u + (1-\pi) r_d > 0$$

Since the terminal wealth W is equal to proceeds from the investment portfolio, we can define W as:

$$W = w(1 + \lambda r_i), \ \ \textit{where } i = \{u, d\}$$

where r_i is a realization of the net risky asset return in the end of the period, and λ is the share of wealth invested in the risky asset.

At first we consider the case of an economy without a and with homogeneous agents. In that scenario the agent solves the following optimization problem:

$$\max_{\lambda} E\left[\frac{W^{1-\gamma}}{1-\gamma}\right], \ \gamma > 0$$

$$s.t \quad W = w(1 + \lambda r_i) - F, \quad \text{for} \quad i = \{u, d\}$$

If the agent decides to invest in the stock market, she faces a fixed cost of stock market participation, F, in the current period. F measures the cost of obtaining information about the stock market and how it functions. The objective function of each agent investing in the risky asset is the following:

$$\max_{\lambda > 0} \frac{\pi (w(1 + \lambda r_u) - F)^{1-\gamma} + (1-\pi)(w(1 + \lambda r_d) - F)^{1-\gamma}}{1-\gamma}$$

The first-order condition for this problem is given by:

$$\pi r_u (w(1 + \lambda r_u) - F)^{-\gamma} + (1-\pi) r_d (w(1 + \lambda r_d) - F)^{-\gamma} = 0$$

Solving for λ, the optimal fraction of the portfolio allocated to risky asset is:

$$\begin{aligned} \lambda^* &= \frac{(1-m)(w-F)}{w(mr_u - r_d)}, \ where \\ m &= \left(\frac{\pi r_u}{(\pi-1)\, r_d}\right)^{-\frac{1}{\gamma}} \end{aligned}$$

Since we assume that $r_d < 0 < r_u$ and $\pi r_u + (1-\pi)\, r_d > 0$, we have that $0 < m < 1$ and as a result $\lambda \geq 0$.[4]

We can see that:

$$\lambda'_w = \frac{F(1-m)}{w^2(mr_u - r_d)} > 0 \tag{1}$$

$$\lambda'_F = -\frac{(1-m)}{w(mr_u - r_d)} < 0 \tag{2}$$

As we can see from (1)) and (2) the risky share increases with initial wealth and decreases with participation costs. In other words, wealthy agents, who can afford to pay fixed stock market participation costs, allocate a higher ratio of their wealth to the stock market than "poor" agents. In addition, when the fixed costs are greater then the initial wealth, $F > w$, the agent does not invest at all.

Now we will introduce a social network to the model. We assume that all agents belong to a social network. If an agent decides to participate in the stock market, this agent becomes an Informed agent. These agents can share information about the stock market with their peers for free. There are two types of agents in the economy: Educated and Non-Educated.[5] The stock market participation of both types of agents depends on their wealth and risk-aversion. The two types of agents differ in their stock market participation costs F. Educated agents have knowledge about the stock market, meaning that the fixed cost for acquiring information is low. In the beginning of the period Non-Educated agents do not have information about the stock market, and their fixed costs are therefore high. Hence, if there is no information sharing between different types of agents Non-Educated agents are less likely to participate at a given level of wealth due to their higher fixed cost.

The fixed cost of participating for Non-Educated agents is assumed to be decreasing in the number of Informed Peers. In other words, the larger is the number of peers that invest in stocks, the smaller is the fixed cost of acquiring information for Non-Educated agents. As a result, a

[4]We assume that investors can borrow money at risk-free rate, so we don't exclude the situataion when $\lambda > 1$.

[5]The term Educated refers here to an agent that is educated about financial markets.

Non-Educated agent i face fixed participation costs equal to $\frac{\theta}{k_i}$, where θ is a constant and k_i is the number of "neighbors" (peers connected in the network) of agent i who participate in stock market.[6] Fixed stock market participation cost for Educated agents are not affected by the number of Informed peers, as they are assumed to already have sufficient information. Educated agents instead face fixed costs of entry to the stock market, θ_P. In summary, the fixed cost for agent i is:

$$F\left(k_i\right)=\begin{cases}\frac{\theta}{k_i} & \text{if } i \text{ is Non-Educated agent}\\ \theta_{Ed} & \text{if } i \text{ is Educated agent}\end{cases} \tag{3}$$

Agent i participates in the stock market only if the certainty-equivalent gain in wealth associated with doing so exceeds her fixed cost. This requires that:

$$\frac{\pi\left(w_i\left(1+\lambda_i^* r_u\right)-F\left(k_i\right)\right)^{1-\gamma}+(1-\pi)\left(w_i\left(1+\lambda_i^* r_d\right)-F\left(k_i\right)\right)^{1-\gamma}}{1-\gamma}-\frac{w_i^{1-\gamma}}{1-\gamma}>0 \tag{4}$$

where w_i defines initial wealth, λ_i^* is the optimal risky share, and $F\left(k_i\right)$ is the participation costs for agent i. If cost $F(k_i)$ is sufficiently high, such that $\left(w_i\left(1+\lambda_i^* r_d\right)-F\left(k_i\right)\right)<0$, then no agent invests in the risky asset.

We assume that agents in the economy belong to a social network. In the model the structure of the network is described by $\{\mathsf{N}, G, W\}$, where N is a set of agents-nodes of power n - the number of individuals in the economy, G is a $n \times n$ matrix describing connections between nodes in the network, and $\mathcal{W}$ is an array of the length n describing the level of the initial wealth allocated to each agent in the network. We assume that the initial wealth for agent i is chosen by a random draw from a distribution with a non-negative lower bound of a support since we assume that initial wealth is non-negative, i.e. $w_i \sim iidF(\cdot)$. We have $\mathcal{W}=\{w_1, ..., w_n\}$.

We assume that there are two types of agents in the economy, i.e. we have two types of nodes in the network as well. We define the type as t, which is equal to 0 if an agent is a Non-Educated, and equal to 1 if the agent is Educated. The type of agent i, t_i is defined as follows:

$$t_i=\begin{cases}0 & \text{if } x_i<\overline{x}\\ 1 & \text{if } x_i>\overline{x}\end{cases} \tag{5}$$

where x_i is a random variable, $x \sim U\{0,1\}$.

[6]Note that when $k_i=0$, participation cost are equal to ∞. Non-Educated do not have direct access to information and remain uninformed about the possibility to participate in the stock market. This is a technical assumption and we can assume instead that this cost is not infinitely high but sufficiently high, which will not affect results.

The level of connectivity between two neighbors (nodes in the network) depend on two factors: the random connectivity parameter, y, and the difference in wealth levels between two neighbors. The social network literature shows that network formation depends on similarity of agents, denoted *homophily* (McPherson, Smith-Lovin, and Cook, 2001). While there are several dimensions of homophily in network that we could use, we chose wealth as an indicator of homophily. This assumption means that the wealth distribution will determine the probability that two agents are connected in theoretical model. We define the difference in wealth levels between two agents i and j as a function $d\left(w_i, w_j\right)$. This function defines "the distance" between two nodes. The closer agents are to each other in wealth, the smaller is the distance measure between them. Parameter y is a random variable uniformly distributed on the interval $[0,1]$, $y \sim U[0,1]$. The connectivity is defined in the following way:

$$g\left(i,j\right)=\begin{cases}0 & \text{if } yc^{-d(w_i,w_j)}<\overline{y}\\ 1 & \text{if } yc^{-d(w_i,w_j)}>\overline{y}\end{cases} \tag{6}$$

If $g\left(i,j\right)=1$, we say that nodes i and j are connected. In the same manner if $g\left(i,j\right)=0$, we say that nodes i and j are disconnected.

In the equation (6) $c>0$ and $\overline{y}$ are model parameters. Parameter c and function $d\left(w_i, w_j\right)$ allow us to establish the correlation between the probability of a connection between agents i and j and difference in the wealth levels $|w_i - w_j|$. If $0<c<1$, we assume that this correlation is negative, so that "poor" agents have a higher probability of connecting (communicating) with rich agents than with "poor". When $c=1$ there is no correlation, and we assume that the probability of a connection between two agents does not depend on the difference in the level of wealth between them. In this paper, we focus on the situation when $c>1$, which corresponds to the situation when agents are more likely to connect with agents who has a similar level of wealth than with agents who are richer or poorer. Parameter $\overline{y}$ is a cutoff parameter for a binomial distribution which corresponds to a network's connectivity.

2.2 Algorithm

In this section we describe the algorithm which allows us to find unique static equilibrium in the network. This equilibrium is the unique static equilibrium, for a given dynamic process of information dissemination with initial conditions. In other words, we can find other static equilibria which contain more active links, but the system cannot reach them through the dynamic information diffusion process.[7]

The network is defined with a matrix of linked nodes and a stack of weights for each node. Each link represents a possibility for information sharing, and each weight corresponds to a wealth level of each node (agent).

Matrix G represents links between agents through which they can communicate and share information. Array W represents the initial wealth of each consumer.

$$G = \{g(i,j), \forall i,j \in \mathbf{N} \text{ such that } g(i,j) = 1 \text{ if i and j are linked, and } g(i,j) = 0 \text{ otherwise}\}$$

$$W = (w_1, .., w_n), w_i \sim F(\cdot)$$

We apply the following dynamic information process algorithm to find an equilibrium. Let's call vector **IP** a vector of all Educated nodes, and vector **P** a vector of all nodes which participate in the stock market (*Informed*). Therefore, the vector of nodes which are Educated but do not participate, is $\mathbf{I} = \mathbf{IP} - \mathbf{P}$.

At step 0 of the algorithm we assume that fixed participation costs for Educated nodes are equal to zero, $\theta_{Ed} = 0$.[8] We have that vector **IP** coincides with a vector of initially Informed nodes, meaning that all Educated agents invest in the risky asset.

At step i: $\mathbf{P_i} = \mathbf{1}_{\{x_j\}_{j=\overline{1,n}}>0}(f(\mathbf{W}) - c(\mathbf{G} \times \mathbf{IP_i}))$, where $\mathbf{1}(\cdot), f(\cdot), c(\cdot)$ are multivariate scalar functions. For each node we solve a stock market participation problem as in (4). If node x decides to participate in the stock market, we mark it as a participant, i.e. $x \in P_i$.

At step $i+1$: $\mathbf{IP_{i+1}} = \mathbf{1}_{\{x_j\}_{j=\overline{1,n}}>0}(G \times P_i + P_i)$. Since the node x is a participant now, the total number of informed neighbors for each node y, such that $g(x,y) = 1$, increases by one.

We repeat the step i. Algorithm stops at step k when $\mathbf{P_k} = \mathbf{P_{k-1}}$.

Theorem 1. *An equilibrium reached through an algorithm described above is unique and does not depend on order in which we treat all nodes.*

Proof. At first, we show that the partition of **N** on participants and non-participants which we get as a result of our algorithm is indeed an equilibrium. It is obvious that any node which is included

[7]By an active link we mean a link through which nodes transmit information relevant to stock market participation. Each node which participate in the stock market we call an active node.

[8]We make this assumption to simplify the simulation process. The same algorithm can be used in the case of $\theta_{Ed} > 0$.

in the set of participants P has positive profit and has to belong to this set (invest in risky asset). Now let us consider the set of nodes which are not included in the set of participants. Here we can consider two subsets: informed nodes I and uninformed nodes U.[9] Every node $i \in I$ appears in the stack of informed nodes after any neighbor $\{j \in P | g(i, j) = 1\}$ who participates, so if the optimal decision at the last treatment was "not participate" it remains the same after the end information diffusion process as the set of informed neighbors of node i remains the same. Now we can consider the node $i \in U$. We assume that this node is uninformed about the possibility to participate in stock market so it has fixed cost which is equal to ∞ and does not want to participate. So we can see that we have indeed an equilibrium.

Secondly, we show that the order in which we treat nodes does not affect the final set of participants P. The first observation that the node which decides to participate at the step i would not change the decision if it was treated the last time at the step $j > i$ as fixed cost does not increase over time. If this node was treated the last time at the step $j < i$ and changed its opinion, that means that from step i to step j new information appears, and step j can not be the last when this node is treated. Any node which is treated the last time at the step i and decides not to participate also cannot change the decision if it is treated at step $j < i$ (following the argument above) and, in addition, since the step j is the last one when this node is treated, from step j to the terminal step no relevant information for this node appears (no one of the its neighbors become informed). Thus, if the node was treated later, it would make the same decision. $\square$

Figure (1) shows an example of a network and the equilibrium that we reach with the algorithm described above. The algorithm progresses from the initial state to the second and final iteration. At the initial state Agent 1 is Educated and finds it beneficial to invest in stocks. As a result, all the agents connected to Agent 1, Agents 3 and 2, have an additional Informed peer in their networks. Agent 2 has one Informed peer in her network, which means that the fixed stock market participation cost for Agent 2 is equal to 6. Since the utility of stock market participation is now positive $(8 - 6 = 2)$, Agent 2 invests in stocks. Furthermore, Agent 3 has 3 peers in her network, but only one Informed peer who participates, Agent 1. Since the fixed cost of 6 for Agent 3 is larger than the benefit of 5 for investing in the stock market, the expected utility of stock investment is negative and Agent 3 does not participate. At this point the dynamic information diffusion process

[9]We include in set I only informed nodes which do not belong to set P

stops. Agents 4 and 5 have no Informed peers in their networks. As a result of the second iteration, we have an equilibrium of 40% (2 agents out of 5) stock market participation.

To further illustrate how social network effects will influence the estimated stock market participation costs, consider estimating the fixed cost for stock market participation with the above network structure and fixed costs. The average cost for stock market investors in the above example is $(0 + 6/1)/2 = 3$, the average for Agent 1 and 2. This cost is only half of the true cost for Agent 2 – the real cost for Non-Educated is half than the real cost. Clearly taking the average effect and neglecting the role of social networks in spreading information can bias estimated effects. This example illustrates why it is important to take into account network effects and agent heterogeneity while estimating stock market participation costs.

As mentioned above, there are other types of static equilibrium that we can not obtain through the dynamic process. In order to illustrate this idea, consider the following simple example. Suppose that we have five agents and network structure as in Figure (2).

In Figure (2) there is another static equilibrium achieved through a different algorithm. Similar to before we have that the Educated Agent 1 participates in the stock market, and therefore Agent 1 shares information with Agent 2 that lowers the cost sufficiently so that Agent 2 also participates. Now, consider Agents 3 and 5. We can see from the figure that Agent 5 will participate if she has one Informed peer, and that Agent 3 will participate if she has 2 informed peers. There is therefore a static equilibrium where Agent 5 participates because Agent 3 participates, and Agent 3 participates because Agent 1 and 5 participate. Obviously this is a static equilibrium, but we can not reach it through the dynamic process. Stock market participation in this equilibrium is 80 percent (4/5), higher than the previous dynamic equilibrium.

We call this equilibrium the most efficient equilibrium, and we provide an algorithm to compute it. By the most effective equilibrium we mean the static equilibrium such that the maximum possible number of nodes participates. We allow all Non-Educated to be initially Informed, meaning that they all participate in the stock market. Then we apply the following algorithm:

We start with original graph G. At first, we create two stacks where we add the computed thresholds $\tilde{k}_i$ for each consumer and a number of neighbors koi. The thresholds represent the number of participating neighbors which each agent needs in order to participate. Secondly, for every node we compute an expected utility of participation based on the agent's private information and cost. We compare number of neighbors with a threshold for every node. Every node for which condition $\tilde{k}_i > k_i$ is satisfied is excluded form the graph. So we pass from graph G_j to G_{j+1} with reduced number of nodes. Finally, we repeat the procedure while $G_{j+1} \subseteq G_j$ and $G_j \nsubseteq G_{j+1}$. The graph which we obtain at the end gives as the equilibrium with the biggest set of nodes.

We can add other components of participants in this algorithm, or increase the size of an original component. We avoid any situation when a Non-Educated node with only private information participates. In other words, the size of each component which contains a Non-Educated node is not less than two.

Theorem 2. *Algorithm 2 finds the unique maximum stable equilibrium.*

Proof. At first we consider the static equilibrium in which all nodes are informed and share information. If under this condition everybody wants to participate, then we have an equilibrium which includes all nodes. If there is a node which does not want to participate even if all its neighbors are informed, then this node can not belong to a set of participants in any equilibrium. As it does not belong to a set of participants, it also can not share any information. So we can exclude this node and all its links without affecting the final outcome as this node and all links are not active in the equilibrium. As a result, we have that each time we apply this procedure we have a reduced game with the same equilibrium. So we apply the same procedure recursively. At some step we can not remove any node because either the set of nodes is empty or when all remained nodes want to participate given that all of them share the information about the stock market. So we have an equilibrium. That equilibrium is characterized by the biggest set of participants as we know that every eliminated node does not belong to an equilibrium set. The equilibrium is unique because the order of elimination does not matter. □

Above we define the upper limit of stock market participation in the economy with a given social network structure. Social network structure defines the maximum number of stock market participants in an economy, as saw above in Figure (2) – the maximum achievable stock market participation is 80% (4 agents out of 5).

3 Simulations

In this section, we test how the theoretical model described above fit the stock market participation observed in the data. First, we use simulations to show how the parameters of the model (number of Educated agents, average wealth, and connectivity) affect stock market participation. We show that low connectivity in the network can explain low participation in the stock market even if the average wealth in the economy is high. At the same time, we can reach an equilibrium with high stock market participation even when the average level of wealth is low if connectivity in the network is high.

The literature has discussed the role of financial literacy/education/employment. Several studies indicate that financial illiteracy is widespread and that individuals lack knowledge of even the most basic economic principles (Lusardi and Mitchell, 2007; Lusardi and Mitchelli, 2007; Lusardi and Mitchell, 2011; Hilgert, Hogarth, and Beverly, 2003).[10]

Figure (3) shows that an increase in the percent of Educated agents increases stock market participation. In the model an increase in the proportion of Educated agents leads to more efficient propagation of information about the stock market, all else equal. Even if the proportion of Educated agents is low, it has a strong impact on the information diffusion.

The main contribution with this paper is that we show that parameters of the social network in the economy, especially connectivity between people, affects the level of stock market participation. In the theoretical model, the connectivity depends on a random element and on parameters of the wealth distribution. The random element is uniformly distributed on the interval from 0 to 1. The two features of the wealth distribution that are important for us are average wealth and Gini Index as a measure of wealth inequality. The probability of being connected in the network in the model depends on how similar individuals are in their wealth. A low Gini Index indicates that wealth inequality is low, which increases the probability that two agents are connected.

Figure 4 show how stock market participation depends on the wealth parameters. The left side graph shows the results for the Gini Index, and the right shows the result for wealth.

These two parameters define the wealth distribution. Low level of social inequality has a positive impact on stock market participation. Furthermore, a higher level of average wealth in the economy

[10] There are concerns that households are not saving enough for the retirement, are accumulating excessive debt, and are not taking advantage of financial innovation (Lusardi and Mitchelli, 2007; Campbell, 2006). Existing studies have also shown that those who are not financially literate are less likely to plan for retirement and to accumulate wealth (Lusardi and Mitchell, 2007; Lusardi and Mitchell, 2011) and are more likely to take up high-interest mortgages (Moore, 2003) or have problems with debt (Lusardi and Tufano, 2009). Finally, Van Rooij, Lusardi, and Alessie, 2011 show that many families shy away from the stock market because they have little knowledge of stocks, the mechanisms of the stock market, and the asset pricing.

implies that agents have more disposable funds to invest. Thus, the participation rate increases when the disposable income grows. In addition, high wealth makes the fixed participation cost relatively lower. In the economy with a social network structure, wealthy Non-Educated agents need to have a relatively low number of participating peers to decrease their participation costs.

We also see that an increase in Gini Index (i.e. a lower wealth inequality) leads to higher stock market participation. By construction in our model, interconnectedness between individual agents (nodes) depends on wealth distribution. We assume that the probability for agents with the similar wealth to be peers is higher than that probability for agents with different wealth levels. If we observe a relatively flat wealth distribution, the probability of agents to be connected is higher. In a low inequality society therefore, each agent potentially has more peers. In other words, it is easier for an agent to get more information from peers and to decrease her stock market participation costs when everyone has similar level of wealth.

Figure (5) shows how stock market participation depends on the expected risk premium and volatility. The result is straightforward - a higher risk premium and lower volatility are associated with a higher participation rate.

One of the main advantages of our theoretical model is that it allows us to be flexible with the parameters of the economy, and allows us to vary several characteristics at once. In the below exercise we show how different parameters can interact. For example, even though the number of educated agents in the economy is high, the stock market participation can be low if connectivity is low, as information does not spread throughout the economy. Figure (6) shows all the equilibria that we can generate in the model for different levels of connectivity and the number of financially educated agents.

In Figure (7) we show possible equilibria that we can reach by manipulating connectivity and average wealth in the economy. We see that a high level of participation is possible even if the average level of wealth in the economy is low. The reason for this is that connectivity in the model is high, which helps spread information and reduces the fixed cost of stock market participation. Conversely, if the average wealth is high, but individuals are not well connected, the stock market participation is limited, since there is no information sharing between wealthy stock market investors and other agents in the economy.

Figure (8) shows equilibria in the model when we manipulate values for the Gini Index and mean wealth. Recall that the connectivity parameter in the model depends on the level of the wealth equality. A more equal wealth distribution leads to a higher connectivity, which increases stock market participation. A very high level of mean wealth itself does not necessarily mean that stock market participation is high, as the wealth can be concentrated among agents who belong to small social networks.

Finally, Figure (9) shows the results of the model simulations for different values of the Gini Index and the number of Professionals. Similarly, to the previous results, we see that since connectivity depends on the Gini Index the level of stock market participation can be low, even if the number of professionals in the economy is high. Conversely, participation can be high if connectivity is high, even if the number of Educated agents is high.

4 Estimation

In the theoretical model, we assume that the stock market participation of an individual depends on the fixed stock market entry costs and on the number of Informed agents in agent's network. In this section, we test our assumptions using Danish administrative data.

4.1 Data Description

We assemble a dataset from the universe of the Danish population focusing on individuals of age 18-65 years old for the period from 2010 to 2013. The dataset contains economic, financial, and demographic information about all individuals, including an identifier for their geographical location. We aggregate the individual-level data to the municipal level, and consider each municipality as individual entity (an economy) in the simulations.[11] We also have data on a smaller geographical unit (a "shire") that we will use later to construct measures of connectivity. There are approximately 3000 shires in Denmark.

The dataset is constructed based on several different administrative registers made available from Statistics Denmark. Individual and family data originate from the official Danish Civil Registration System. These records include the personal identification number (CPR), gender, date of birth. In addition to providing individual characteristics, such as age, gender, and marital status, these data enable us to identify municipalities where agents are registered. The dataset provides unique identification across individuals, households, generations, and time.

[11] There are 98 municipalities in Denmark.

Income, wealth, and portfolio holdings are obtained from the official records at the Danish Tax and Customs Administration (SKAT). This dataset contains personal income and wealth information by CPR numbers on the Danish population. SKAT receives this information directly from the relevant sources; financial institutions supply information to SKAT on their customers deposits and security investments. Employers similarly supply statements of wages paid to their employees. Through Statistics Denmark, we obtain access to this information from 2010 to 2013.

Educational records are obtained from the Danish Ministry of Education. All completed (formal and informal) education levels, as well as identification numbers of schools and universities attended by individuals, are registered on a yearly basis and made available through Statistics Denmark. We use these data to measure individuals education level and whether they have a financial education. We also use this data to construct networks, which is described more below.

The Danish Labor Ministry provides Statistics Denmark with information on workplace, sector, and the length of the employment for each individual. All the records are linked to individual CPR numbers. That helps us to match all the dataset and to collect all necessary information for estimation.

Our parameter of interest is the connectivity in the network, i.e. number of informed/stock market participating peers that an individual has in her social network. First, we assume that individuals from the same household are connected. Second, the literature shows that social networks are characterized by homophily (McPherson, Smith-Lovin, and Cook, 2001). We use several dimensions of homophily among agents to reconstruct the social network. In particular, we use information about the workplace, university, and age. We first construct an indicator for age differences as:

$$I_{ageCohortWP_{ij}} = \begin{cases} 1 & \text{if} \quad |age_i - age_j| < 4 \\ 0 & otherwise \end{cases}$$

Colleagues and ex-classmates indicator functions are then constructed in the following way:

$$I_{CS_{ij}} = \begin{cases} 1 & \text{if individuals } i \text{ and } j \text{ study at the same school,} \\ 0 & otherwise \end{cases} \tag{7}$$

$$I_{CW_{ij}} = \begin{cases} 1 & \begin{array}{c}\text{if individuals } i \text{ and } j \text{ studied together or} \\ \text{work at the same place}\end{array} \\ 0 & otherwise \end{cases}$$

Using workplace, university and age, we construct proxies for connectivity:

$$ConnectivityWorkplace_{ij} = \left\{ \; I_{CW_{ij}} * I_{AgeCohort_{ij}} \; \right\} \tag{8}$$

$$ConnectivityUniversity_{ij} = \left\{ \; I_{CS_{ij}} * I_{AgeCohort_{ij}} \; \right\} \tag{9}$$

where $I_{CW_{ij}}$ is an indicator function which is equal to 1 if individuals i and j work at the same place; $I_{CS_{ij}}$ is an indicator function which is equal to 1 if individuals i and j graduated from the same educational institution. $I_{ageCohortWP_{ij}}$ is an indicator function that shows if the absolute value of the difference in age between individuals i and j is less than 4 years. We condition connectivity on age to make sure that individuals belong to the same cohort at the university and the same workplace.

The general level of connectivity is determined by both workplace and university. We define the general level of connectivity (or density) of the social network in the municipality m as (10):

$$Density_m = 1 - \left(1 - \frac{\sum_{j=1}^{N_m} ConnectivityWorkplace_{ij}}{N_m}\right)\left(1 - \frac{\sum_{j=1}^{N_m} ConnectivityUniversity_{ij}}{N_m}\right) \tag{10}$$

We use a 5% representative sample of the population for each municipality in the simulations. We ensure that the sample is representative by preserving the same income and age distributions in samples as in the general population in each municipality. In addition, we ensure that the ratio of financially literate and the ratio of employed agents is the same. Table (1) describes the average number of connections in each type of network and shows that the number of peers is stable across time within each network. Table (2) provide the summary statistics of the data for all years for all individuals and for Participants and Nonparticipants separately.

Table (2) shows that individuals who invest in stocks (Participants) on average are wealthier, older, studied for longer, and have a lower number of kids under the age of 18 years. Participants are also more likely to be married and more likely to have studied economics or/and finance. They have higher connectivity and have more stock market investors among their peers. These results are stable over time.

4.2 Results

Equation 1 shows that the optimal risky share in the investment portfolio is decreasing in the (unobservable) fixed costs. We therefore turn towards estimating how participation is affected by the number of peers in the network. The prediction from the model and equation (3) is that informed peers share knowledge with their uninformed peers, and that this reduces their fixed cost of participation. Given that the risk premium is positive and similar across individuals, the prediction from the model is that the optimal risky share is increasing in the number of peers investing in risky assets. We therefore estimate the following probability model:[12]

$$\begin{aligned} Prob[\lambda^*_{i,t} > 0] = {} & \alpha SMP_{i,t-1} + \beta Network_{i,t} + \gamma logWealth_{i,t} \\ & + \delta_1 Occupation_{i,t} + \delta_2 LogIncome_{i,t} + \delta_3 Age_{i,t} + \delta_4 AgeSqr_{i,t} \\ & + \delta_4 Numofkids_{i,t} + \delta_5 Education_{i,t} + \delta_6 MaritalStatus_{i,t} \\ & + \mu Year_t + \epsilon_{i,t}, \end{aligned} \tag{11}$$

where for individual i in year t, $\lambda^*_{i,t}$ represents the risky share in the investment portfolio, $SMP_{i,t-1}$ is a dummy equal to 1 if individual invested in stocks in the previous period and 0 otherwise, $Network_{i,t}$ defines the number of peers participating in the stock market, $logWealth_{i,t}$ is the logarithm of total wealth, and $Year_t$ is a year fixed effect. We use demographic characteristics to control for individual specific effects. Specifically, $Occupation_{i,t}$ is a dummy variable equal to 1 if individual i is employed in financial sector in year t; $LogIncome_{i,t}$ represents the logarithm of the labor income of individual i in year t; $Age_{i,t}$ defines the age of individual i in year t; $Numofkids_{i,t}$ equals to the number of kids younger than 18 years of individual i in year t; $Education_{i,t}$ represents the length of schooling of individual i in year t; and finally $MaritalStatus_{i,t}$ is a dummy equal to 1 if individual i is married in year t and 0 otherwise.

Results from the estimation of (11) are presented in Table (3). First, lagged stock market participation, $SMP_{i,t-1}$, is positive and strongly significant and provides evidence for some inertia in portfolio choice. Second, connectivity has positive and significant effect on propensity to invest in stocks, thus providing support for the predictions of the model. The other coefficients have the correct sign and correspond well to the previous literature – wealth, the level of education and financial literacy have strong positive effects on the probability of investing in risky assets.

[12] As a benchmark we use the model described in Vissing-Jorgensen (2002) and Khorunzhina (2013).

In the theoretical model we assume that financially educated individuals are not affected by the number of informed peers in their network, as they have already acquired the necessary information. To test this hypothesis, we include an interaction term between financial literacy and the number of informed peers in column (2). The interaction of connectivity with financial literacy variable is significant and negative, but not economically significant. This result is consistent with our assumption that for financially educated people fixed stock market participation costs are independent from the connectivity in the social network.

Tables 4 and 5 describe the results for alternative measures of connectivity: number of informed peers for individuals with network, Ratios of informed peers among all the peers, density of the social network, and density of informed social network. The results are consistent with the results of the previous tables.

4.3 Model Comparison

In this section we provide results of simulations. To begin, we compute parameters of the population density, mean income, the Gini coefficient, and percentage of professionals for each municipality. We then create a random sample of n individuals for each municipality with corresponding parameters, and then apply an algorithm to find an equilibrium, while varying the parameter of stock market cost for non-professionals.

We have to estimate whether our model improves the prediction of stock market participation compared to a model using only income, Gini Index, and number of financially educated individuals as predictors. In the previous section we considered which parameters affect individual choice of stock market participation. However, we argue that parameters of social interactions should also affect the level of SMP in particular social group (municipalities, regions, countries).

In order to estimate performance of linear regression model we do the following procedure. We take all observations for years 2010, 2011, 2012 and 2013 for 98 municipalities. So we have 396 observations in total. We sample them in random order and divide into two groups. The first group is a regression sample, which includes 299 observations. The second group is a test sample, which includes 97 observations. We estimate parameters of linear regression on regression sample:

$$SMP_i = \beta_0 + \beta_1 \text{mean income}_i + \beta_2 \text{gini}_i + \beta_3 \text{financial employment}_i + \beta_4 \text{density of connections}_i + \varepsilon_i \tag{12}$$

The results of the linear regression analysis are presented in Table 6.

After we obtain the regression coefficients, we use them to forecast stock market participation on the test sample and compare the results with what we obtain through simulations. Note that we do not estimate any parameters for our simulations except for participation cost θ.[13] Instead we directly apply our algorithm to the parameters from the data, and calibrate for the appropriate choice of θ which we then use for the whole sample. Figure 10 illustrates the results of this exercise. As we can see, the simulated model predicts SMP better than regression analysis. Considering the fact that both models use the same parameters as input, we can conclude that the underlying structure of information diffusion indeed affects stock market participation.

Comparison results are presented in figure 10. On the left we see the comparison of the predicted by regression stock market participation. On the right we see the prediction results from the simulations. We see that simulated results match actual stock market participation better than the regression model.

In the theoretical model we also propose the concept of the Maximum Stable equilibrium. Using the Algorithm 2 we define the maximum stock market participation in municipalities in the simulated sample. Figure 12 presents simulation results for unique static equilibrium and for the maximum stable equilibrium. As we can see on average the maximum stock market participation ratio is greater that the equilibrium stock market participation. We also notice that for some municipalities the maximum participation ratio matches the equilibrium participation.

In the theoretical model the maximum stock market participation coincides with equilibrium stock market participation if the density of the network is higher. If we look at the figure 13, we see that the density is decreasing among the municipalities in simulation sample. The difference between maximum participation and equilibrium participation is negatively correlated to the density of the social network. The regression results, described in the table 7, prove that hypothesis. Higher density in the social network induces information diffusion and increases the equilibrium stock market participation to the point when it reaches the maximum participation level.

[13] In Figure 11 we illustrate the results of simulations for 2010 with different choice of θ.

5 Conclusion

In this paper we introduce a theoretical model describing the risk-taking decisions of agents in an economy with a social network. By simulating the model, we show that changes in the wealth distribution, stock market performance and financial education level cause changes in the rate of stock market participation. In the model, the parameters of wealth distribution determine the level of connectivity between individuals. In particular, "similarity" in wealth between individuals directly affect the probability of agents to be connected in the network. We show that the higher is wealth inequality in the economy, the lower is stock market participation.

Based on the results of the theoretical model, connectivity in the economy is an important determinant of the level of stock market participation. We provide also empirical results showing that the propensity to invest in stock for an individual positively correlated with the number of informed investors in her social network.

Simulation results show that the model with network predicts the stock market participation better than the regression model. Also the results reveal that the equilibrium participation converges to the maximum stock market participation when the density of the social network increasing.

References

Ang, Andrew, Geert Bekaert, and Jun Liu (2005). "Why stocks may disappoint". In: *Journal of Financial Economics* 76.3, pp. 471–508.

Arrow, Kenneth Joseph (1965). *Aspects of the theory of risk-bearing*. Yrjö Jahnssonin Säätiö.

Brav, Alon, George M Constantinides, and Christopher C Geczy (2002). *Asset pricing with heterogeneous consumers and limited participation: Empirical evidence*. Tech. rep. National bureau of economic research.

Brown, Jeffrey R et al. (2008). "Neighbors matter: Causal community effects and stock market participation". In: *The Journal of Finance* 63.3, pp. 1509–1531.

Bursztyn, Leonardo et al. (2014). "Understanding mechanisms underlying peer effects: Evidence from a field experiment on financial decisions". In: *Econometrica* 82.4, pp. 1273–1301.

Calvet, Laurent E, John Y Campbell, and Paolo Sodini (2007). "Down or out: Assessing the welfare costs of household investment mistakes". In: *Journal of political Economy* 115.5, pp. 707–747.

Campbell, John Y (2006). "Household finance". In: *The Journal of Finance* 61.4, pp. 1553–1604.

Cao, H Henry, Tan Wang, and Harold H Zhang (2005). "Model uncertainty, limited market participation, and asset prices". In: *Review of Financial Studies* 18.4, pp. 1219–1251.

Changwony, Frederick K, Kevin Campbell, and Isaac T Tabner (2014). "Social engagement and stock market participation*". In: *Review of Finance*, rft059.

Christiansen, Charlotte, Juanna Schröter Joensen, and Jesper Rangvid (2008). "Are economists more likely to hold stocks?" In: *Review of Finance* 12.3, pp. 465–496.

Dow, James and Sergio Ribeiro da Costa Werlang (1992). "Uncertainty aversion, risk aversion, and the optimal choice of portfolio". In: *Econometrica: Journal of the Econometric Society*, pp. 197–204.

Epstein, Larry G and Martin Schneider (2007). "Learning under ambiguity". In: *The Review of Economic Studies* 74.4, pp. 1275–1303.

Fama, Eugene F and Kenneth R French (2002). "The equity premium". In: *The Journal of Finance* 57.2, pp. 637–659.

Guiso, Luigi and Tullio Jappelli (2005). "Awareness and stock market participation". In: *Review of Finance* 9.4, pp. 537–567.

Guiso, Luigi, Paola Sapienza, and Luigi Zingales (2008). "Trusting the stock market". In: *the Journal of Finance* 63.6, pp. 2557–2600.

Haliassos, Michael and Carol C Bertaut (1995). "Why do so few hold stocks?" In: *the economic Journal*, pp. 1110–1129.

Heaton, John and Deborah Lucas (2000). "Portfolio choice and asset prices: The importance of entrepreneurial risk". In: *The journal of finance* 55.3, pp. 1163–1198.

Hilgert, Marianne A, Jeanne M Hogarth, and Sondra G Beverly (2003). "Household financial management: The connection between knowledge and behavior". In: *Fed. Res. Bull.* 89, p. 309.

Hong, Harrison, Jeffrey D Kubik, and Jeremy C Stein (2004). "Social interaction and stock-market participation". In: *The journal of finance* 59.1, pp. 137–163.

Kaustia, Markku and Samuli Knüpfer (2012). "Peer performance and stock market entry". In: *Journal of Financial Economics* 104.2, pp. 321–338.

Khorunzhina, Natalia (2013). "Structural estimation of stock market participation costs". In: *Journal of Economic Dynamics and Control* 37.12, pp. 2928–2942.

Lusardi, Annamaria and Olivia S Mitchell (2007). "Baby boomer retirement security: The roles of planning, financial literacy, and housing wealth". In: *Journal of monetary Economics* 54.1, pp. 205–224.

— (2011). *Financial literacy and planning: Implications for retirement wellbeing.* Tech. rep. National Bureau of Economic Research.

Lusardi, Annamaria and Olivia Mitchelli (2007). "Financial literacy and retirement preparedness: Evidence and implications for financial education". In: *Business economics* 42.1, pp. 35–44.

Lusardi, Annamaria and Peter Tufano (2009). *Debt literacy, financial experiences, and overindebtedness.* Tech. rep. National Bureau of Economic Research.

Mankiw, N Gregory and Stephen P Zeldes (1991). "The consumption of stockholders and nonstockholders". In: *Journal of financial Economics* 29.1, pp. 97–112.

McPherson, Miller, Lynn Smith-Lovin, and James M Cook (2001). "Birds of a feather: Homophily in social networks". In: *Annual review of sociology*, pp. 415–444.

Mehra, Rajnish and Edward C Prescott (1985). "The equity premium: A puzzle". In: *Journal of monetary Economics* 15.2, pp. 145–161.

Moore, Danna L (2003). *Survey of financial literacy in Washington State: Knowledge, behavior, attitudes, and experiences.* Washington State Department of Financial Institutions.

Patacchini, Eleonora and Edoardo Rainone (2014). “The word on banking-social ties, trust, and the adoption of financial products”. In: *Tech. Rep., Einaudi Institute for Economics and Finance (EIEF)*.

Van Rooij, Maarten, Annamaria Lusardi, and Rob Alessie (2011). “Financial literacy and stock market participation”. In: *Journal of Financial Economics* 101.2, pp. 449–472.

Vissing-Jorgensen, Annette (2002). *Limited asset market participation and the elasticity of intertemporal substitution*. Tech. rep. National Bureau of Economic Research.

Zenou, Yves (2012). “Networks in economics”. In:

6 Figures

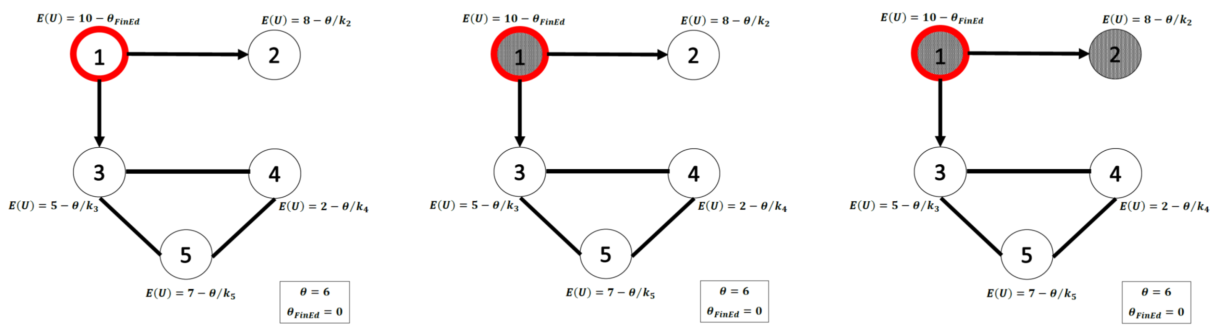


Figure 1a: Initial State

Figure 1b: First Iteration

Figure 1c: Second Iteration

Figure 1: Unique Static Equilibrium

θ defines the fixed stock market particiaption costs for Non-Educated agents, θ_{FinEd} - fixed stock market particiaption costs for Educated agents, k_i is the number of Informed peers in an agent's network; $E(U)$ - expected utility of investing in stocks. Circles outlined by red define Educated agent and outlined by black define Non-Educated.

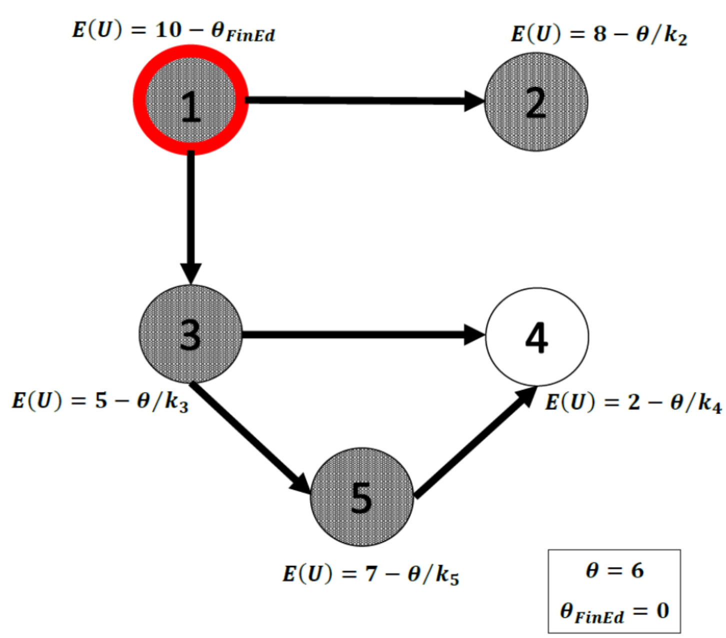


Figure 2: Example: The Most Effective Equilibrium

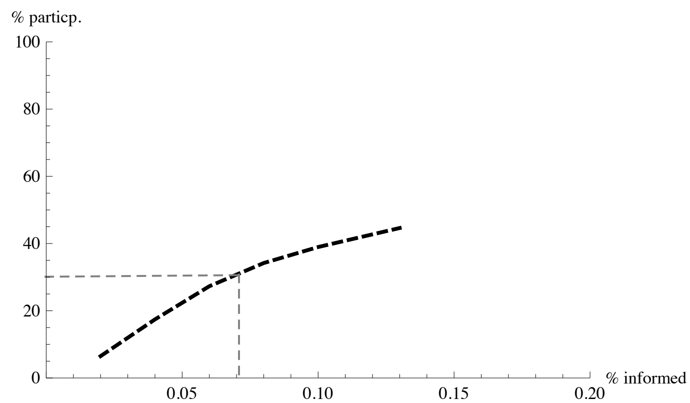


Figure 3: Stock Market Participation & Level of Financial Education

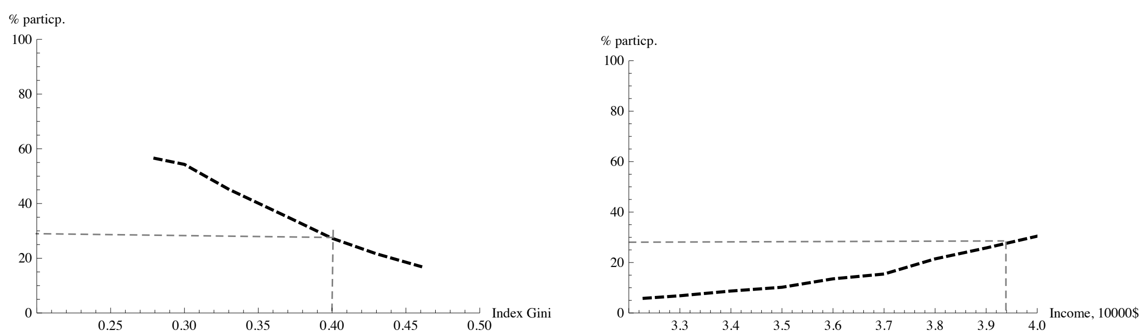


Figure 4: Stock Market Participation & Wealth

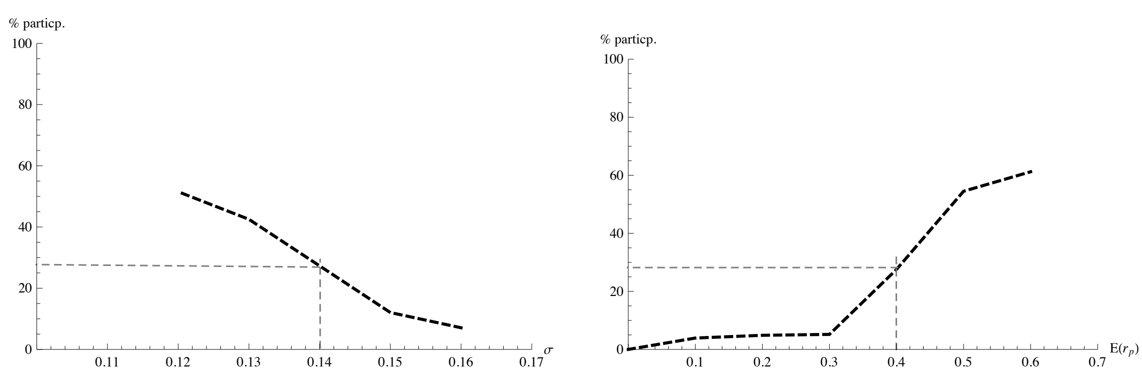


Figure 5: Stock Market Participation & Stock Market Parameters

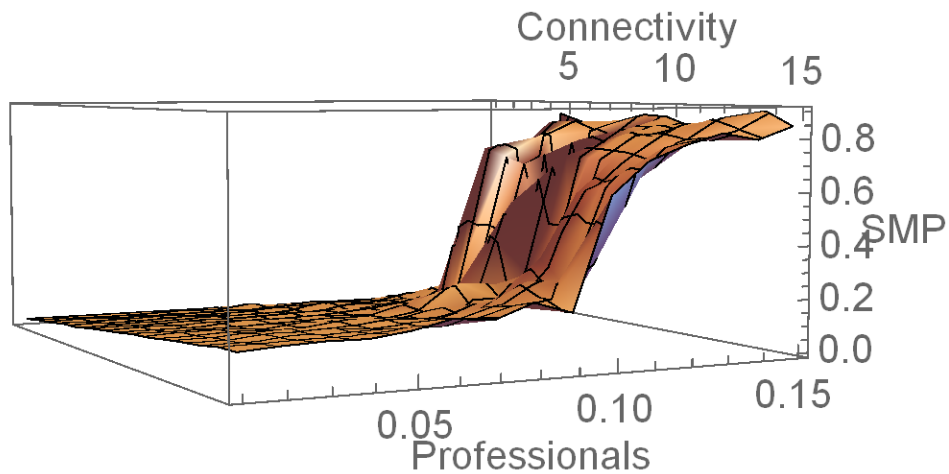


Figure 6: Stock Market Participation : Level of Financial Education & Connectivity

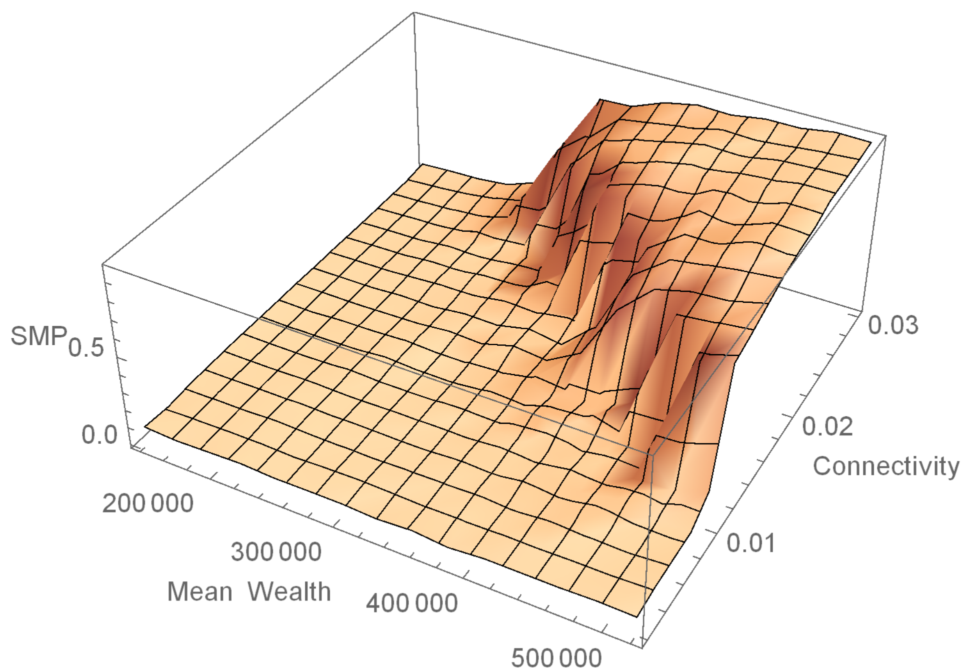


Figure 7: Stock Market Participation : Mean Wealth & Connectivity

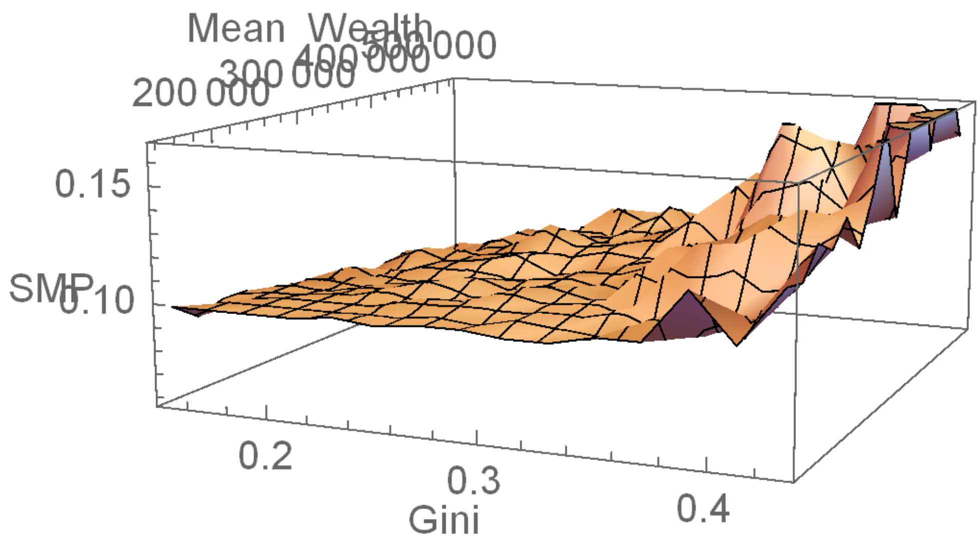


Figure 8: Stock Market Participation : Gini Index & Mena Wealth

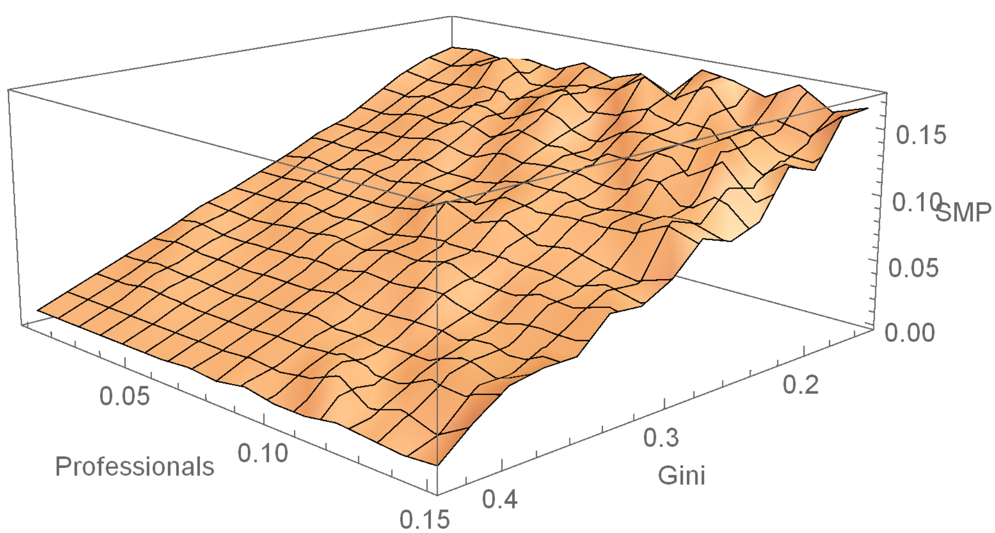
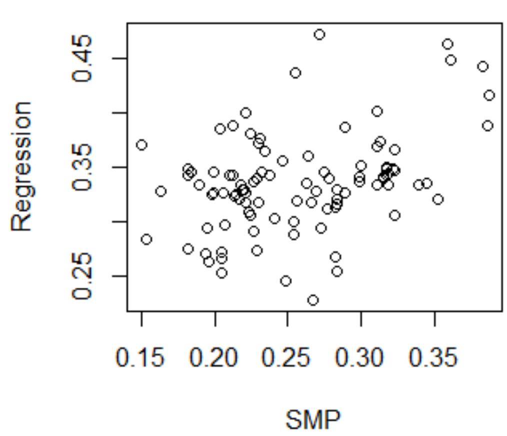
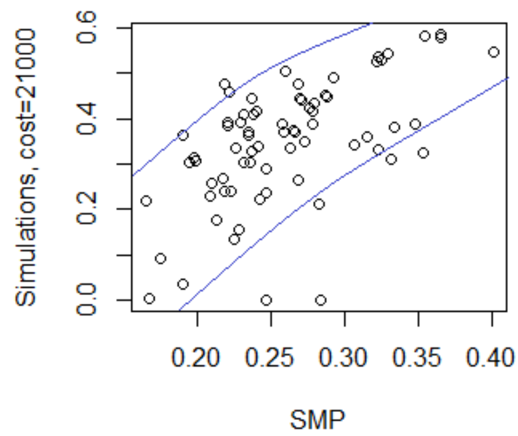


Figure 9: Stock Market Participation : Gini Index & Number of Educated Agents



(a) Regression



(b) Simulations

Figure 10: Stock Market Participation Forecast

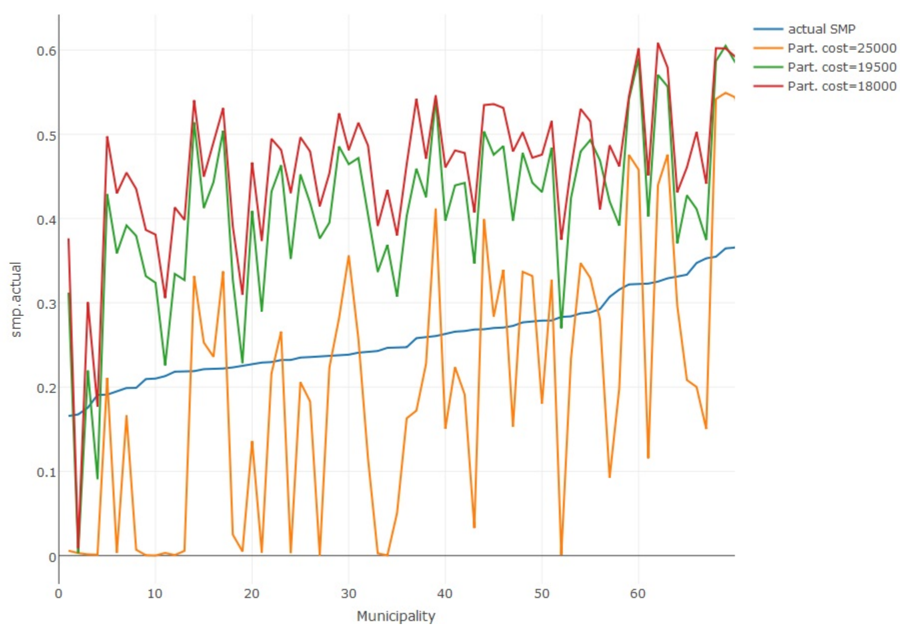


Figure 11: Stock Market Participation : Results of Simulations for 2010

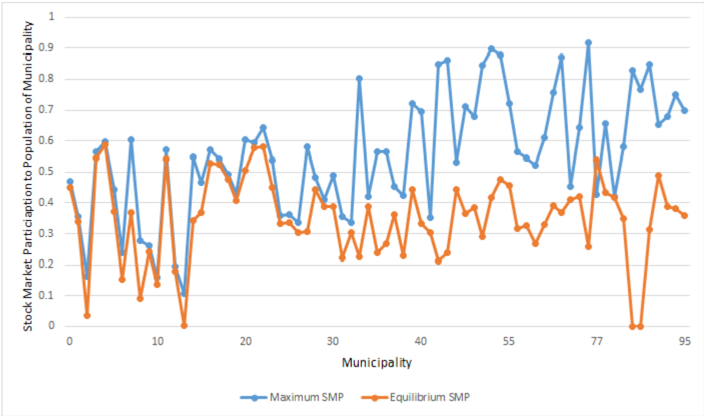


Figure 12: Maximum SMP and Equilibrium SMP, 2010

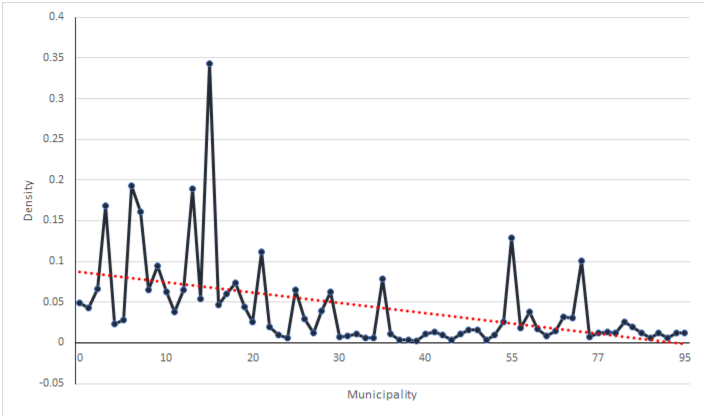


Figure 13: Maximum SMP and Equilibrium SMP, 2010

7 Tables

Table 1: Number of connections in the network

Year	Workplace	University	Shire
2010	1.58	14.82	29.83
2011	1.54	14.95	29.00
2012	1.48	15.16	29.49
2013	1.43	15.76	30.14

Notes: Average number of connections for individuals. Number of connections for workplaces, university and shire calculated according to equations (8), (9) and (**??**) respectively.

Table 2: Summary Statistics For All Households

	ALL	PARTICIPANTS	NON-PARTICIPANTS
SMP	0.25	1.00	0.00
	(0.44)	(0.00)	(0.00)
SMP in (t-1)	0.26	0.95	0.02
	(0.44)	(0.23)	(0.14)
Peers	746.83	911.29	690.82
	(1,361.75)	(1,479.48)	(1,314.58)
Peers - Stock Investor	81.31	320.12	0.00
	(1,361.75)	(1,479.48)	(1,314.58)
Log of Labor Income	12.49	12.63	12.45
	(1.14)	(1.19)	(1.12)
Log of Total Wealth	6.59	9.34	5.66
	(6.21)	(5.94)	(6.02)
Age	41.09	43.97	40.10
	(12.29)	(12.22)	(12.16)
# of kids under 18	0.79	0.75	0.80
	(1.02)	(1.00)	(1.03)
Length of schooling	12.95	13.50	12.77
	(2.84)	(2.83)	(2.81)
Marital Status	0.51	0.57	0.49
	(0.50)	(0.50)	(0.50)
Occupation: Financially Employed	0.00	0.00	0.00
	(0.02)	(0.03)	(0.02)
Financial literacy	0.04	0.09	0.03
	(0.20)	(0.28)	(0.16)
# of Colleagues - Stock Investors	27.22	32.20	25.53
	(86.09)	(85.94)	(86.07)
# of Ex-Classmates - Stock Investors	721.87	881.82	667.41
	(1,353.76)	(1,470.42)	(1,397.22)
Observations	10,095,679	2,564,407	7,531,263

Notes: Summary statistics for all households. Standard deviations in parentheses. SMP is a dummy for stock market participation. Occupation measures number of individuals employed in the financial sector. # of colleagues and # of ex-classmate is not conditional on being within the same age bracket.

Table 3: OLS: The effect of Informed Peers

	(1)	(2)
SMP in $(t-1)$	0.886***	0.885***
	(5648.92)	(5617.26)
# of Informed Peers	0.000119***	0.000131***
	(510.52)	(488.37)
Log of Labor Income	0.000659***	0.000624***
	(11.97)	(11.36)
Financially Literate	-0.0124***	0.00363***
	(-39.83)	(10.17)
Age	-0.00224***	-0.00220***
	(-53.69)	(-52.93)
Age Squared	0.0000325***	0.0000319***
	(64.77)	(63.63)
# of Kids under 18	-0.000222**	-0.000224**
	(-3.01)	(-3.05)
Length of Education	-0.00134***	-0.00153***
	(-57.63)	(-65.46)
Marital Status	0.000141	0.000127
	(0.96)	(0.87)
# Informed Peers × Financial Literacy		-0.0000437***
		(-91.78)
YEAR FE	YES	YES
REGIONAL FE	YES	YES
Observations	7,166,988	7,166,988

Notes: * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$. T-statistics is described in the brackets. SMP is a dummy for stock market participation. Number of Informed peers is equal to the number of coworkers an classmates from the same age cohort, who invest in stocks.

Table 4: OLS: Th effect of Networks on the Stock Market Participation

	I. Informed Peers of Individuals With Network		II.Ratio of Informed Peers	
	(1)	(2)	(1)	(2)
SMP in $(t-1)$	0.883***	0.882***	0.654***	0.645***
	(5 391.49)	(5 364.79)	(3 105.09)	(3 050.70)
Informed Peers	0.000122***	0.000134***		
	(515.41)	(490.83)		
Ratio of Informed Peers			0.919***	0.976***
			(1 592.40)	(1 611.08)
Log of Labor Income	0.000635***	0.000600***	-0.000148**	-0.000111*
	(10.75)	(10.18)	(-2.90)	(-2.17)
Financial Education	-0.0176***	-0.00102**	-0.0514***	0.0172***
	(-53.65)	(-2.68)	(-181.92)	(46.9)
Age	-0.00219***	-0.00214***	0.00136***	0.00154***
	(-49.35)	(-48.27)	(35.18)	(40.18)
Age Squared	0.0000319***	0.0000312***	-0.0000267***	-0.0000295***
	(59.16)	(57.76)	(-56.94)	(-63.09)
Kids under 18	-0.000167*	-0.000175*	-0.00133***	-0.00123***
	(-2.18)	(-2.29)	(-20.02)	(-18.58)
Length of Education	-0.00158***	-0.00179***	-0.00102***	-0.00157***
	(-63.42)	(-71.28)	(-47.93)	(-74.32)
Marital Status	0.0000189	0.0000276	-0.00240***	-0.00240***
	(0.12)	(0.18)	(-18.10)	(-18.18)
Connectivity Parameter × Financial Literacy		-0.0000413***		-0.314***
		(-85.15)		(-292.44)
YEAR FE	YES	YES	YES	YES
REGIONAL FE	YES	YES	YES	YES
Observations	6,663,414	6,663,414	6,663,414	6,663,414

Notes: * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$. t-statistics in brackets. Model I estimates the effect of informed peers on the stock market participation of individuals (classmates and coworkers from the same age-cohort). Model II estimates the effect of the ratio of informed peers among all the peers on the stock market participation.

Table 5: OLS: Th effect of Networks on the Stock Market Participation

	III. Network		IV. Inf. Network	
	(1)	(2)	(1)	(2)
SMP in $(t-1)$	0.919***	0.919***	0.909***	0.909***
	(6 317.00)	(6 316.21)	(5 862.64)	(5 861.7)
Density of the Network	0.00110*	0.00145**		
	(2.32)	(3.02)		
Density of Informed Network			1.041***	1.053***
			(272.03)	(248.54)
Log of Labor Income	0.00154***	0.00154***	0.00133***	0.00133***
	(27.51)	(27.53)	(22.21)	(22.21)
Financial Education	0.0158***	0.0166***	0.00575***	0.00663***
	(50.46)	(45.34)	(17.5)	(18.75)
Age	-0.00172***	-0.00172***	-0.00175***	-0.00175***
	(-40.64)	(-40.62)	(-38.93)	(-38.87)
Age Squared	0.0000256***	0.0000255***	0.0000258***	0.0000258***
	(50.06)	(50.03)	(47.21)	(47.13)
Kids under 18	-0.000103	-0.0000974	-0.000467***	-0.000463***
	(-1.38)	(-1.30)	(-6.01)	(-5.96)
Length of Education	0.00134***	0.00134***	0.000600***	0.000593***
	(57.67)	(57.78)	(24.16)	(23.88)
Marital Status	-0.000268	-0.000263	-0.000770***	-0.000766***
	(-1.79)	(-1.76)	(-4.96)	(-4.94)
Connectivity Parameter × Financial Literacy		-0.0117***		-0.0622***
		(-4.43)		(-6.73)
YEAR FE	YES	YES	YES	YES
REGIONAL FE	YES	YES	YES	YES
Observations	6,663,414	6,663,414	6,663,414	6,663,414

Notes: * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$. t-stat in brackets. Model III estimates the effect of the density of the social network (# of peers over total # of citizens in the municipality). Model IV estimates the effect of the density of informed network.

Table 6: Stock Market Particiaption in Municipalities, 2010

Parameter	Coefficient	$Pr(> F)$	Signif.
Mean Income	2.255597e-07	5.556e-09	***
Gini	8.641221e-01	5.481e-10	***
# of Financially Educated Individuals	-3.045539e+01	2.462e-07	***
Density	1.689983e+00	0.09995	

Table 7: Density and Equilibriums

	(1)
	MaximumSMP
Density	0.721***
	(10.12)
Constant	0.405***
	(19.47)
Observations	71

t statistics in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Dependent variable is the difference between maximum stock market participation and the equilibrium participation.

Social Investment Styles

Steffen Andersen
Copenhagen Business School

Olga Balakina
Bocconi University

Kim Peijnenburg
HEC, Paris

Abstract

In this paper we construct a model of investment choice in the presence of peer effect and empirically test whether the portfolio choices of individuals are influenced by the choices of their peers using data for the full Danish population from 2000 to 2013. We combine tax records with matched employer-employee data to construct networks based on workplaces, geography, and family connections, and test whether networks influence households portfolio choices, and which network has the largest effect on stock market participation decisions. We use an instrumental variable strategy to show that co-worker networks exert the largest influence on the probability to participate in the stock market, followed by family networks. Geographical networks have a limited influence. Furthermore, we see that the relative influence of each network depends on the characteristics of the network and the characteristics of the household. Networks with higher levels of education, income, and age have a larger impact on choices. Women are more influenced by their family network, and men by their co-worker network. Furthermore, greater similarities between the household and the average network characteristics result in a higher influence of that network.

1 Introduction

The time has come to move beyond behavioral finance to social finance, which studies the structure of social interactions, how financial ideas spread and evolve, and how social processes affect financial outcomes.

David Hirschleifer, 2015

Contrary to most rational models in economics, individual decisions are rarely made in isolation. Every day, our decisions are influenced by our friends and family, by our co-workers and neighbors, and by others whose actions dictate the norms in society. A growing field within finance and economics, *social finance*, tries to understand how social networks affect individual behavior and how these insights can be incorporated into models of economic decisions. The primary purpose of this paper is to investigate the role of social networks for investment decisions both theoretically and empirically.

We begin by constructing a theoretical model to show how the decision to invest in risky assets depends on the investment choices of peers. In the model an individuals' expected return depends both on their own beliefs and on the beliefs of peers. When an individual observes that their peers invest in the stock market, she correctly infers that the peer expects the risk premium to be positive. The average stock market participation within an individual's social network therefore corresponds to the average belief within the network.

We use data on the full population of Danish households from 2000 to 2012 to explore the properties of social interactions and analyze the impact of different types of peer networks on household's investment decisions. Specifically, we identify networks of colleagues, family, and neighbors using high quality administrative data, and estimate how each network affects the probability of either entering or exiting the stock market. To account for time-invariant characteristics that affect the probability of investing in the stock market we examine the decision to enter or exit the stock market. Furthermore, it is well known in the literature on peer effects that endogeneity concerns are prevalent. To account for possible endogeneity we use an instrumental variable strategy based on third degree peers, following Bramoullé and Fortin (2009). We estimate whether social networks affect changes in stock market participation using two separate methodologies. First, we use the change in stock market participation as the dependent variable and estimate using both Ordinary Least Squares (OLS) and an instrumental variable strategy based on third degree peers. Second, we use a Multinomial Logistic Model (MLOGIT) to estimate separate effects for the decision to enter and to exit. With the later strategy we use the instrumented decision to exit in a first stage to account for endogeneity.

In our first contribution, we show that each group has an impact on stock market participation when estimated in separate regressions. These results confirm the finding of the previous literature (citations). However, when we pool all networks into one regression, only co-worker networks have a significant effect on stock market participation. These results hold in the OLS, IV and MLOGIT estimation. Furthermore, we use the time variation in stock market returns to examine whether the effects of social networks vary over time. Our sample covers a period in Denmark with both stable returns (2000-2003), a housing and stock market boom period 2004-2008), a collapse in the housing and stock market (2008) and a subsequent recovery. Nevertheless, we find that the co-worker effect is stable across time, and that other networks sporadically have a significant impact on stock market participation.

Our paper contributes to the literature on the determinants of stock market participation, a literature that has received considerable attention from researchers in both finance and economics. Only half of U.S. households invest in stock market and even smaller fraction, fewer than 20 percent, hold stocks or mutual funds outside of retirement plans.[1] This lack of stock investments among retail investors is known as the "limited stock market participation puz-

[1] In Denmark the stock market participation outside of pension funds is approximately 27 percent.

zle", which highlights that while portfolio theory predicts a positive risk share if the expected risk premium is positive (Mehra and Prescott, 1985), the actual participation is substantially lower. Different explanations for the limited stock market participation puzzle have been discussed in literature. The following explanations of limited stock market have been proposed: high fixed costs of stock market participation Vissing-Jorgensen, 2002; Mankiw and Zeldes, 1991; Heaton and Lucas, 2000, lack of stock market awareness Hong, Kubik, and Stein, 2004; Guiso and Jappelli, 2005; Brown and Taylor, 2010, non-standard preferences like ambiguity aversion Dow and Costa Werlang, 1992; Ang, Bekaert, and Liu, 2005; Cao, Wang, and Zhang, 2005; Epstein and Schneider, 2007, education deficits Campbell, 2006; Calvet, Campbell, and Sodini, 2009; Christiansen, Joensen, and Rangvid, 2008; Van Rooij, Lusardi, and Alessie, 2011, lack of trust Guiso, Sapienza, and Zingales, 2008. In addition, our paper can contribute toward us understanding the types of mistakes that households when they invest, such as portfolio under-diversification, disposition effect, home bias, and other. On the macro level the literature investigates the events, such as asset bubbles and herding in investments. A non-exclusive list of reasons for "excessive" stock investment include: learning from peers choices can lead to spread of investment mistakes, and as a result to asset bubbles Bikhchandani and Sharma, 2000; Chari and Kehoe, 2004, or peer's possession of an asset makes individuals to overinvest in that asset Abel, 1990; Gali, 1994; Campbell and Cochrane, 1999; DeMarzo, Kaniel, and Kremer, 2004; DeMarzo, Kaniel, and Kremer, 2008. Understanding herding behavior and how they create asset bubbles clearly requires an understanding how investment decisions spread throughout social networks.

Our paper also contributes to empirical and experimental literature discussing peer effects on investment decisions (for example, Kaustia and Knüpfer (2012), Changwony, Campbell, and Tabner (2014), and Bursztyn et al. (2014)). Peer groups are defined in a wide range of ways, starting from a general sociability of individuals (Hong, Kubik, and Stein (2004)) to specific groups of peers such as co-workers and family Li (2014). Specifically, the literature has defined groups of peers as: socially engaged individuals (Changwony, Campbell, and Tabner (2014)), co-workers (Duflo and Saez (2002)), neighbors (Kaustia and Knüpfer (2012) and Brown et al. (2008)), family (Li (2014)). All the above papers assume the presence of a single network (peer group). Our paper contributes by showing that when we include three separate networks, each group has a significant impact. However, when all three networks are included in one regression, only co-worker networks have a significant effect on stock market participation. This is consistent with Patacchini and Rainone (2014), who shows that within networks there are weak and strong ties, and that not all social contacts are equally important. To our knowledge, this is the first paper to examine the effects of different types of networks simultaneously.

The paper also contributes to the literature discussing the econometrics of networks. The literature on identification and estimation of social network models has progressed significantly recently (see Blume and Ioannides (2011) and Durlauf and Ioannides, 2010 for recent surveys). We base our work on the seminal paper by Bramoullé and Fortin, 2009, who generalize the linear-in-means model of Manski (1993) to a general local-average social network model, where the endogenous effect or peer effect is represented by the average outcome of the peers. Our contribution to this literature is to empirically estimate the model with social network effects on stock market participation while taking into account heterogeneity in the network.

The rest of the paper unfolds as follows. We discuss the theory behind the model of peer effect on stock market participation in Section 2. The empirical model and estimation strategy of the econometric network model is detailed in Section 3. Our data and descriptive statistics are described in Section 4. The empirical results of heterogeneous peer effects on stock market participation are provided in Section 5.

2 Theoretical Background

In this section we discuss the theoretical basis of heterogeneous peer effects on household level stock market participation. We propose a model that relates investment behavior of a household with the investment decisions of its peers. We assume that peers can be heterogeneous by dividing individual social network in several groups.

There are two components that combined create a theory of peer effect on stock market participation: portfolio allocation theory and social learning theory. From portfolio theory we have that the risk share depends on the distribution of the risk premium in the market and on the level of risk aversion of the individual (Merton, 1969):

$$\lambda = \frac{\mathbb{E}[R]}{\gamma \sigma_R^2} \tag{1}$$

where λ stands for the share of financial wealth invested in risky assets, R represents the risk premium on the market, $\mathbb{E}[R]$ and σ_R^2 are mean and variance of the risk premium, and γ is a coefficient of relative risk-aversion. In Mehra and Prescott (1985) the risk share depends on the preference parameter γ and the distribution of the risk premium.

Let us consider a situation without peer effects and with a more general definition of individual preferences and of the distribution of the risk premium.[2] Consider investor i's decision to invest in a risky asset. The asset return is given by x, with probability density function $f(x)$, and investor i's utility is $u_i(x) = u(x)$ for all i. Each investor receives information about asset's return - a signal, s_i, coming from a single distribution, with probability density function $g(s_i)$. We assume that when investing in isolation, investor i does not take into consideration any investor j $(i \neq j)$. Thus, investor i is willing to invest if and only if:

$$\int u(x) f(x|s_i) dx \geq \bar{u} \tag{2}$$

[2]In this Section we use the same notation as in Bursztyn et al., 2014.

where $f(x|s_i)$ is the conditional density, and $\bar{u}$ denotes the outside option for the investor. There exists a unique threshold $\bar{s}_1$ such that for any $s_i \geq \bar{s}_1$ investor i is willing to invest.[3] Denote the decision to buy the asset made by investor i by $b_i = 0, 1$. Hence, for an investor i making a decision to invest in isolation:

$$b_i = 1 \Leftrightarrow s_i \geq \bar{s}_1 \tag{3}$$

The individual i invests in a risky asset, $b_i = 1$, if she expects a risk premium, s_i, is higher than the threshold, $\bar{s}_1$. Now let's assume that instead of making his investment choice in isolation, before making his own decision investor i observes the investment decision of investor j which is given by b_j. Assume that investor j made his choice $b_j = 1$ in isolation and hence his decision rule is given by Equation (3). Thus, when investor i observes $b_j = 1$ she correctly infers that $s_j \geq \bar{s}_1$ and she is willing to invest if and only if:

$$\int u(x) f(x|s_i, s_j \geq \bar{s}_1) dx \geq \bar{u} \tag{4}$$

Equation (4) shows that the investor i updates her beliefs about the risky asset return based on her own signal and on the information that her peer, investor j, received. We have two building blocks for our model: optimal investment decision and the effect of social learning on that decision. We are interested in the decision to invest in the risky asset. We have:

$$b_i = \begin{cases} 1, \ if \ \mathbb{E}[R]_i > 0 \\ 0, \quad otherwise \end{cases} \tag{5}$$

where b_i represents individual stock market participation, $\mathbb{E}[R]_i$ is individual belief about the expected value of the risk premium. A positive share of financial wealth invested in stocks

[3]Given that the conditional density satisfies the monotone likelihood ratio property and given mild monotonicity assumptions on the utility function $u(x)$ of the investor.

implies that the household stock market participation is equal to 1, and it is equal to 0 if the risky share is zero. Furthermore, we have that the coefficient of relative risk aversion is greater than zero, $\gamma > 0$, and variance of the risk premium is non-negative, $\sigma_R^2 > 0$. The sign of the value of risky share depends on the sign of the expected risk premium, $\mathbb{E}[R]$.

Now assume that individual i observes the decisions to invest by her peers. We assume that there are three types of peers in a social network of a household. We call them *co-workers*, *neighbors*, *family*, based on the origin of the social connection.[4] All peers in a type of network have the same preferences. For each peer the following proposition applies:

RULE 1. *If peer i has a belief that the expected risk premium is positive, then her risk share is positive, i.e. she invests in stocks*

Since all the peers in the type of the social network are the same under the central limit theorem the average participation rate among peers in a type of the network corresponds to a belief among peers that expected premium is positive.

$$\frac{\sum_{k=1}^{N_j} b_{j,k}}{N_j} \xrightarrow{a.s.} Pr(\mathbb{E}[R_j] > 0),$$

where $j = CW, N, F$ is the network identifier and N_j represents the number of peers in the network j.

In other words, with social learning the individual stock market participation is defined in the following way:

$$b_i = \begin{cases} 1, \; if \; \mathbb{E}[R \mid \Pr(\mathbb{E}[R_{CW}] > 0), \Pr(\mathbb{E}[R_N] > 0), \Pr(\mathbb{E}[R_F] > 0)]_i > 0 \\ 0, \quad otherwise \end{cases} \tag{6}$$

where $Pr(\mathbb{E}[R_{CW}] > 0)$, $Pr(\mathbb{E}[R_N] > 0)$, $Pr(\mathbb{E}[R_F] > 0)$ represent beliefs of co-workers, neighbors, and family of individual i, respectively, about the sign of the expected risk pre-

[4]Networks are discussed in more details in the following Section

mium. Probabilities (as discussed above) can be identified through the average rate of participation among co-workers, neighbors, and family members, respectively. Model (6) describes the investment behavior of peers affects the individual stock market participation.

We see that individual i update her belief about the expected risk premium based on her own prior and on the information from her peers. We thereby show that the beliefs of peers affect individual decision to invest in stocks.

3 Empirical Model and Estimation Strategy

To estimate the effect of peers investment behavior we create the following empirical model:

$$b_{i,t} = \alpha + \beta_1 \mathbb{E}[RP_{i,t}] + \beta_2 \overline{b_{i,t}^{CW}} + \beta_3 \overline{b_{i,t}^{Neigh}} + \beta_4 \overline{b_{i,t}^{F}} + \eta_i + \gamma X_{i,t} + \delta Y_t + u_{i,t} \tag{7}$$

In equation (7), $\mathbb{E}[RP_{i,t}]$ stands for individual i personal belief about the expected risk premium, $\overline{b_{i,t}^{CW}}$, $\overline{b_{i,t}^{Neigh}}$, and $\overline{b_{i,t}^{F}}$ define average stock market participation in coworker, neighbor, and family networks, respectively, η_i represents time invariant characteristics of the household, such as risk-aversion, $X_{i,t}$ stands for individual i characteristics in period t, Y_t represents time fixed effects, and $u_{i,t}$ is an *iid* with normal distribution, $\mathcal{N}(0, \sigma_u^2)$. We are interested in estimating β_1, β_2, and β_3. Signs and the size of those coefficients allow us to understand how peers choices affect individual stock market participation.

There are several issues with the model described in equation (7). The first issue is that η_i is unobservable. If η_i is correlated with average participation in the networks, our estimation results will be biased. To resolve this issue, we use a first differenced model to remove any time-invariant effects.

The second issue is that individual beliefs are also unobservable, and that individual beliefs can be subject to the same shocks as beliefs of peers in the social network. To resolve this issue, we use instrumental variable approach. Both issues are discussed in detail in the following section. As a result, we estimate the following model: In equation (7), $\mathbb{E}[RP_{i,t}]$ stands for individual i personal belief about the expected risk premium, $\overline{b_{i,t}^{CW}}$, $\overline{b_{i,t}^{Neigh}}$, and $\overline{b_{i,t}^{F}}$ define average stock market participation in coworker, neighbor, and family networks, respectively, η_i represents time invariant characteristics of the household, such as risk-aversion, $X_{i,t}$ stands for individual i characteristics in period t, Y_t represents time fixed effects, and $u_{i,t}$ is an iid with normal distribution, $\mathcal{N}(0, \sigma_u^2)$. We are interested in estimating β_1, β_2, and β_3. Signs and the size of those coefficients allow us to understand how peers choices affect individual stock market participation.

There are several issues with the model described in equation (7). The first issue is that η_i is unobservable. If η_i is correlated with average participation in the networks, our estimation results will be biased. To resolve this issue, we use a first differenced model to remove any time-invariant effects.

The second issue is that individual beliefs are also unobservable, and that individual beliefs can be subject to the same shocks as beliefs of peers in the social network. To resolve this issue, we use instrumental variable approach. Both issues are discussed in detail in the following section. As a result, we estimate the following model:

$$\triangle b_{i,t} = \alpha + \beta_1 \triangle \mathbb{E}[RP_{i,t}] + \beta_2 \triangle \overline{b_{i,t}^{CW}} + \beta_3 \triangle \overline{b_{i,t}^{Neigh}} + \beta_4 \triangle \overline{b_{i,t}^{F}} + \gamma X_{it} + \delta Y_t + u_{it} \tag{8}$$

In equation (8) $\triangle b_{i,t}$ stands for changes in individual stock market participation. Specifically:

$$\triangle b_{i,t} = \begin{cases} 1, & \text{if individual} \quad i \quad \text{enters the stock market in year} \quad t \\ 0, & \text{if individual} \quad i \quad \text{does not change her stock market participation in year} \quad t \\ -1, & \text{if individual} \quad i \quad \text{exits the stock market in year} \quad t \end{cases} \tag{9}$$

3.1 The network model

Consider a population of n households[5] partitioned into $\bar{r}$ networks. For the n_r households in the Nth network, their connections with each other are represented by an $n_r \times n_r$ adjacency matrix $\boldsymbol{G_r} = [g_{ij,r}]$ where $g_{ij,r} = 1$ if households i and j are related in the network and $g_{ij,r} = 0$ otherwise. Let $\boldsymbol{G_r^*} = [g^*_{ij,r}]$ be the row-normalized $\boldsymbol{G_r}$ such that $g^*_{ij,r} = g_{ij,r}/\sum_{k=1}^{n_r} g_{ik,r}$ (Patacchini and Rainone, 2014). The financial activity of household i in network r, stock market participation $R_{i,r}$, is given by

$$\triangle b_{i,r} = \phi \triangle \sum_{j=1}^{n_r} g_{ij,r} b_{j,r} + \sum_{k=1}^{p} x_{ik,r}\beta_k + \sum_{k=1}^{p}\Big(\sum_{j=1}^{n_r} g^*_{ij,r} x_{ik,r}\Big)\delta_k + \eta_r + \epsilon_{i,r}, \tag{10}$$

In this model, $\triangle b_{i,r}$ represents the change in the stock market participation of the household i. The aggregate change in stock investments of i's direct contacts, $\triangle \sum_{j=1}^{n_r} g_{ij,r} b_{j,r}$, represents the *endogenous effect*, wherein a household's choice may depend on those of her contacts about their investment activity. $x_{ik,r}$ indicates the k'th exogenous variable accounting for observable differences in individual characteristics (e.g. gender, education, income, family background, etc.) of members of households. $\sum_{j=1}^{n_r} g^*_{ij,r} x_{ik,r}$ is the average value of the exogenous variables over i's direct contacts with its coefficient δ_k representing *the contextual effect*, wherein a household's stock market participation may depend on the exogenous characteristics of her contacts. η_r is a network specific parameter representing *the correlated effect*, wherein households in the same group tend to behave similarly because they face a similar environment (for example, colleagues at work, neighbors in the neighborhood). $\epsilon_{i,r}$ is an i.i.d. error term with zero mean and finite variance σ^2.

Model (10) can be extended to the case of heterogeneous peer effects. If each "ego-network" (i.e. the social contacts of a specific agent, Patacchini and Rainone (2014)) can be split into three different types - colleagues, neighbors, and family - then Model (10) becomes

[5] Since information about stock holdings is provided on household, we conduct our analysis with households and not individuals

$$\begin{aligned}
\triangle b_{i,r} = \phi^{CW} \triangle \sum_{j=1}^{n_r} g_{ij,r}^{CW} b_{j,r} + \frac{1}{g_{i,r}^{CW}} \sum_{k=1}^{p} g_{i,r}^{CW} x_{ik,r} \delta^{CW} \\
+ \phi^{Neigh} \triangle \sum_{j=1}^{n_r} g_{ij,r}{}^{Neigh} b_{j,r} + \frac{1}{g_{i,r}^{Neigh}} \sum_{k=1}^{p} g_{i,r}^{Neigh} x_{ik,r} \delta^{Neigh} \\
+ \phi^{F} \triangle \sum_{j=1}^{n_r} g_{ij,r}{}^{F} b_{j,r} + \frac{1}{g_{i,r}^{F}} \sum_{k=1}^{p} g_{i,r}^{F} x_{ik,r} \delta^{F} \\
+ x_{i,r}{}' \beta + \eta_r + \epsilon_{i,r},
\end{aligned} \tag{11}$$

where $g_{i,r}^{CW} = \sum_{j=1}^{n} g_{ij,r}^{CW}$, $g_{i,r}^{Neigh} = \sum_{j=1}^{n} g_{ij,r}^{Neigh}$, and $g_{i,r}^{F} = \sum_{j=1}^{n} g_{ij,r}^{F}$ are the total number of colleagues, neighbors, and family members each household i has in her social network r. In Model (11) ϕ_{CW}, ϕ_{Neigh}, and ϕ_F represent *the endogenous effects* (i.e. the effect of colleagues, neighbors, and family members investment behavior on one's own stock investment respectively) while δ_{CW}, δ_{Neigh}, and δ_F capture the impact of the exogenous characteristics of the peers, which are allowed to have a varying effect by peer-type.

3.2 Identification and estimation

Our goal is to show that different types of peers in the social network can have heterogeneous effects on stock market participation. While identifying peers and the effects of their decisions on household's stock market participation, we face several identification issues. A number of papers have dealt with the identification and estimation of peer effects with network data (Bramoullé and Fortin, 2009; Liu and Lee, 2010; Calvó-Armengol, Patacchini, and Zenou, 2009; Lin, 2010; Kelejian and Prucha, 2010)). In particular, identification of peer effects in a linear-in-means model is difficult because peers may have similar levels of stock market participation due to: (a) contextual effects, (b) endogenous effects, and (c) correlated effects. In our specific application these three effects could be described as follows: (a) households with highly educated peers may have different investment behavior than those with mostly low-educated peers; (b) there may be genuine peer influences, i.e., investment behavior of households changes in response to the investment behavior of peers; (c) risk-taking behavior of all households in the network is similar, or was affected by the same unobservable shock, such as liquidity shock (De Giorgi, Frederiksen, and Pistaferri, 2015).

We solve these econometric issues by extending the network approach idea of Bramoullé and Fortin (2009) and De Giorgi, Pellizzari, and Redaelli (2010), which rests on the existence of intransitive triads, i.e., friends of friends who are not friends themselves. Note first, however, that contextual effects by controlling for average characteristics of peer in each type of a network. Second, we address the issue of correlated effects with *instrumented peers*, i.e. we instrument each direct peer with a third-degree peer. The key fact is that peers still belong to the same network as a household of interest, but they are *"far enough"* in the network to escape for selection into the network.

To illustrate how we solve the identification problem, let's start from a simple example of a transitive triad (De Giorgi, Frederiksen, and Pistaferri, 2015). The economy consists of three single households 1, 2, and 3. The most general model is one in which changes in stock market participation of a household i (i = 1; 2; 3) depends on her own exogenous characteristics X_i, and on the exogenous characteristics and changes in stock market participation of the other two households. As in Manski (1993), this model is not identified. To see the type of identification strategy we follow, assume now that the households in our example represent the simplest form of an intransitive triad, i.e., household 1's investment behavior is influenced by household 2, who in turn is influenced by household 3, who in turn behaves atomistically. Assume also that household 3 is subject to an exogenous shock T3. Hence, the restricted form is the following:

$$
\begin{cases}
\triangle b_1 = \theta \triangle b_2 + \gamma X_2 \delta X_1 + \epsilon_1 \\
\triangle b_2 = \theta \triangle b_3 + \gamma X_3 \delta X_2 + \epsilon_2 \\
\triangle b_3 = \delta X_3 + \rho T_3 + \epsilon_3
\end{cases}
\tag{12}
$$

The reduced form of this system is:

$$
\begin{cases}
\triangle b_1 = \theta(\gamma + \theta\delta) X_3 + \theta^2 \rho T_3 + (\gamma + \theta\delta) X_2 + \delta X_1 + u_1 \\
\triangle b_2 = (\gamma + \theta\delta) X_3 + \theta \rho T_3 + \delta X_2 + u_2 \\
\triangle b_3 = \delta X_3 + \rho T_3 + u_3
\end{cases}
\tag{13}
$$

In equations (12) and (13) $\triangle b_i$ represents the change in stock market participation of household i, X_i defines household i characteristics (age, education, occupation of family members, and number of kids for the household), T_3 represents an exogenous shock (firm shock) that household 3 is subject to, ϵ_i is iid with $\mathcal{N}(0, \sigma_\epsilon^2)$ and $\mathbb{E}[\epsilon_i, \epsilon_j] \neq 0$ with $i \neq j$, and u_i is iid with $\mathcal{N}(0, \sigma_u^2)$ $\mathbb{E}[u_i, u_j] \neq 0$ with $i \neq j$.

The system (13) is triangular, and therefore it is easy to see that as long as $(\gamma + \theta\delta) \neq 0$ and $\rho \neq 0$, we can recover all the structural parameters from the reduced form. Identification comes from two sources. First, the exogenous shock to household 3 (T_3) and the characteristics of household 3 (X_3) can be used as an instrument for $\triangle b_2$ in household 1's change of stock market participation equation (in network language, distance-3 peers' exogenous shocks and characteristics are valid instruments). This is because T_3, or alternatively X_3, affect $\triangle b_2$ due to contextual effects in household 2's change in stock market participation equation (12), but it has no direct effect on household 1's change in stock market participation since 1 and 3 are not directly connected as we can see from equation (12).

In our model the dependent variable can take three values: Exit, No Change and Entrance. For discrete outcome variables we can estimate multinomial logistic regression models. To estimate the effect and to account for contextual effects we use a two stage procedure. In the first stage we estimate the value of the change in the coworkers' stock market participation using Demographic IVs, Firm IVs (exogenous instruments), and endogenous instruments from the IV model described before. At the second stage, we estimate a mlogit model where the change in households' participation is the dependent variable, and the predicted value of the change in coworkers' participation from the first stage are used as explanatory variables (including other controls as well).

4 Data and Networks

In the following section we describe the data that we use to estimate the impact of different types of networks, and describe how we identify the networks.

4.1 Data

We assemble a dataset of economic, financial, and personal information for the universe of the Danish population focusing on married couples for the time period from 2000 to 2013. To simplify the network identification, we make sure that each spouse is employed in different firms. We use spousal relations and employment at the same place as ties in the co-workers network. We use information about parental ties to identify family members. Information about addresses allows us to identify neighbor networks.

The dataset is constructed based on several different administrative registers made available from Statistics Denmark. Individual and family data originate from the official Danish Civil Registration System. These records include the personal identification number (CPR), gender, date of birth, CPR numbers of family members (parents, children, and thus siblings), and their marital histories (number of marriages, divorces, and widowhoods). In addition to providing individual characteristics, such as age, gender, and marital status, these data enable us to identify all parents. The dataset provides unique identification across individuals, households, generations, and time.

Income, wealth, and portfolio holdings are from the official records at the Danish Tax and Customs Administration (SKAT). This dataset contains personal income and wealth information by CPR numbers on the Danish population. SKAT receives this information directly from the relevant sources, as financial institutions supply information to SKAT about their customers deposits and holdings of security investments. Employers similarly supply statements of wages paid to their employees. Through Statistics Denmark, we obtain access to this information from 2000 to 2013. For the same period, we have also information about stock and mutual fund holdings.

Educational records are from the Danish Ministry of Education. All completed (formal and informal) education levels are registered on a yearly basis and made available through Statistics Denmark. We use these data to measure education levels.

4.2 Colleague Network

To identify co-workers we use a firm identifier. We assume that all the people that are working in the same firm in the same year are colleagues. Based on our IV strategy, we use the investment behavior and characteristics of a third-degree peers to estimate contextual and endogenous effects. We assume that spousal and work relationship are each equal to one degree of distance in the network. In our context, we call individual j a peer of individual i if individual j is a colleague of a spouse of a colleague of individual i.

Figure 1 provides an example of how we identify instruments for peers in the co-workers network. In this example, we have five households with both spouses and four different firms. Both spouses are employed and they work in different firms.[6] We also have four different firms. In this setting we can identify three households that are going to be third-distance nodes in the same network of colleagues. For example, in household 1 the wife works with

[6]Data and restrictions on the sample are discussed further in the Subsection 3.4

husband from household 2. *Wife* 2 in her turn works at the same firm as wife from the household 3. *Wife* 1 and *Wife* 3 are colleagues of spouses of colleagues for each other. Thus, *Wife* 1 is a co-worker instrument for *Wife* 3 and Wife 3 is an instrument for a colleague for *Wife* 1. If we apply the same logic to the rest of individuals in our example, we can see that *Wife* 2 and *Husband* 4 are instruments, as well as *Husband* 3 and *Husband* 5 are instrumented colleagues for each other.

4.3 Neighbour Network

We define neighbors as households living in the same geographical location as the household of interest. To identify *Neighbors* network we use a geographical identifier, a shire.[7] The map of shires in Denmark is presented in Figure 3.

The proximity in the networks corresponds to the degrees of distance in the network. We assume that one "distance" is equal to one for households within the same shire. Thus third-distance neighbors are households from the third furthest shire. By identifying neighbors that live in the third furthest shires we take into account the selection of people into the area (correlated effects). At the same time, we preserve the network structure.

4.4 Family Network

We analyze two types of family network: *Close family* and *Family instruments*. One of the problems associated with the identification of the peer effect in the network is correlated effects. The behavior of individuals can be similar because there is a self-selection into a network. If we assume that there is no self-selection into the family, we can use close family members as peers. However, this seems unlikely and we therefore we apply our instrumental strategy. To define the third-degree distance relatives, we use spousal, parental, and links

[7] A shire is a small geographical location based on historical church district. There are approximately 3,000 shires in Denmark.

between siblings.

Figure 4 shows that the close family includes siblings and children both of husband and wife.[8] For the wife in the household of interest one of the examples of a third-distance relative would be her brother's wife's parents.

4.5 Descriptive Statistics

Table 1, Table 2, and Table 3 show descriptive statistics for the sample and the co-workers, neighbors, and family networks respectively for years from 2000 till 2013. The close family is represented by children and siblings, and as expected the average age of close relatives is lower than the average age of members of households in interest.

5 Results

This section presents the results when we estimate the effect that different types of networks have on changes in stock market participation. We begin by presenting the results of each network separately, and proceed to include all networks within the same regression model. We first present results from an OLS and an IV estimation first, where the dependent variable is the change in stock market participation. Second, we estimate a Multinomial Logit model that allows us to estimate the effect that networks have on the separate decision to enter or exit the stock market.

5.0.1 Coworkers Network

The most commonly used group of peers is coworkers. The literature show that investment decisions of coworkers have strong effect on individual decisions to invest (e.g. Duflo and

[8] Figure 4 shows only wife's side of the tree. For the analysis we take into account both husband and wife sides of the family

Saez, 2002). We therefore start by estimating the following model of household stock market participation:

$$\triangle b_{i,t} = \alpha_0 + \alpha_1 \triangle \overline{b_{i,t}^{CW}} + \alpha_2 \overline{X_{i,t}^{CW}} + \delta_1 X_{i,t}^{Wife} + \delta_2 X_{i,t}^{Husb} + \gamma_1 YearFE_t + \gamma_2 RegionFE_i + \epsilon_{i,t}, \tag{14}$$

where for every year t and every household $i, \triangle b_{i,t}$ defines change in stock market participation, $\triangle \overline{b_{i,t}^{CW}}$ represents change in average stock market participation among household i's coworkers, $\overline{X_{i,t}^{CW}}$ represents the average characteristics of coworkers. $YearFE_t$ is year fixed effect, $RegionFE_i$ is region fixed effect. The results are presented in the Table 4. Column 1 presents the results from an OLS estimation of equation (16), and column 2 presents the results from an instrumental variable strategy where we instrument the change in the average stock market participation among household i's co-workers using the strategy defined above.

As we can see from the table, a change in the average participation of coworkers increases the probability that households enter the stock market. According to the theoretical model, a change in average participation of coworkers corresponds to a shift in the belief of coworkers about the expected risk premium. After observing that shift a household updates its beliefs about the risk premium and decides to enter the stock market. If the average participation of coworkers increases, it corresponds to an increase in the expected risk premium, thus the household is more likely to enter the stock market. If the average participation of coworkers decreases, it corresponds to the decrease of the expected risk premium, as a result the household is less likely to enter the stock market.

The results from the MLOGIT estimation are presented in Table 5, and shows that coworkers' investment decisions significantly increase the probability that households enter the stock market. However, the decision to exit is unaffected by coworkers investment decisions.

5.0.2 Family Network

The literature shows that the likelihood of entering the stock market is higher for people whose family members had entered the stock market (e.g. Li (2014)). We estimate the following model of family effect on household's stock market participation:

$$\triangle b_{i,t} = \alpha_0 + \alpha_1 \triangle \overline{b^F_{i,t}} + \alpha_2 \overline{X^F_{i,t}} + \delta_1 X^{Wife}_{i,t} + \delta_2 X^{Husb}_{i,t} + \gamma_1 YearFE_t + \gamma_2 RegionFE_i + \epsilon_{i,t}, \tag{15}$$

$\triangle \overline{b^F_{i,t}}$ represents change in average stock market participation among household's family members and $\overline{X^F_{i,t}}$ represents the average characteristics of family members. The results for the linear regression model and two stage instrumental variable model are presented in the Table 6.

The results of the OLS model show that change in stock market participation of family members has positive and significant effect on the individual stock market participation. However, when we use instrumental variables the effect disappears. this result can be explained by the fact that family members are subject to unobservable shocks. these shocks create the correlation between stock market participation of a household and its family members.

The MLOGIT model results are provided in the Table 7. To estimate the model, we are using the same approach as in the case of coworkers network. The only exception is that we use only Length of Education of family members as exogenous instrument. Results in Table 7 show that change in stock market participation of family members affect significantly the probability that the household enters the stock market. If the stock market participation of family members increases, the household is more likely to enter the stock market. We don't see any effect on the exit from the stock market.

5.0.3 Neighbors Network

Neighbors are the most discussed peer group in the literature. Starting from Hong, Kubik, and Stein (2004), who discusses how the sociability and the intensity of communication with neighbors increases the probability of stock market participation, the literature shows that investment decisions of neighbors have strong effect on individual decisions to invest. Kaustia and Knüpfer (2012) show that improvement of neighbors' portfolio outcomes increases the number of individuals in the area entering the stock market. In addition, Brown et al. (2008) show that average stock market participation decisions of one's community has an effect on an individual's decision of whether to own stocks.

We estimate the following model of the effect of neighbors on household's stock market participation:

$$\triangle b_{i,t} = \alpha_0 + \alpha_1 \overline{\triangle b_{i,t}^{Neigh}} + \alpha_2 \overline{X_{i,t}^{Neigh}} + \delta_1 X_{i,t}^{Wife} + \delta_2 X_{i,t}^{Husb} + \gamma_1 YearFE_t + \gamma_2 RegionFE_i + \epsilon_{i,t}, \tag{16}$$

where for every year t and every household i $\overline{\triangle b_{i,t}^{Neigh}}$ represents change in average stock market participation among household's neighbors, and $\overline{X_{i,t}^{Neigh}}$ represents the average characteristics of neighbors. The results of the estimation are presented in the Table 8

The results show that average stock market participation of neighbors has positive and significant effect on household's stock market participation, both in OLS and Instrumental variable model.

The results for the multinomial logistic model are presented in the Table 9. From the MLOGIT model we see that average participation of neighbors has significant and positive effect on the entrance probability of the household. The exit probability stays unaffected.

5.0.4 All Social Networks

The previous results suggest that individually each type of network have a significant effect on the household's stock market participation decision, consistent with what the previous literature has found. The key question, however, is whether which network is the most important to the household's decision. To understand and compare the effects of different types of peer groups we pool all types of social networks together and estimate the following model:

$$\begin{aligned}\triangle b_{i,t} = \alpha_0 \\ &+ \alpha_1 \triangle \overline{b_{i,t}^{CW}} + \beta_1 \overline{X_{i,t}^{CW}} \\ &+ \alpha_2 \triangle \overline{b_{i,t}^{Neigh}} + \beta_2 \overline{X_{i,t}^{Neigh}} \\ &+ \alpha_3 \triangle \overline{b_{i,t}^{Fam}} + \beta_3 \overline{X_{i,t}^{Fam}} \\ &+ \delta_1 X_{i,t}^{Wife} + \delta_2 X_{i,t}^{Husb} + \gamma_1 YearFE_t + \gamma_2 RegionFE_i + \epsilon_{i,t},\end{aligned} \tag{17}$$

The first model that we estimate is the dynamic model describing the effect of the social network on the individual stock market participation. Our main interest lies in estimating endogenous effects by regressing individual stock market participation on average stock market participation of three types of networks. The coefficients of interest are α_1, α_2, and α_3. As before, we evaluate contextual effects by controlling for average characteristics of the network (β_1, β_2, β_3). We assume that the effect of each network is constant over time, and later confirm this in the data. We control for the year fixed effects and regional fixed effects.

The results of the estimation are represented in the Table Table 10. The endogenous effect is measured by the coefficients corresponding to the average participation in networks. In the linear model the average stock market participation of all networks have significant effect on the probability of household to invest in stocks. The degree of the effect is different, with Neighbors having the largest effect and Family the lowest. If we turn to results for Exit and Entrance, we can see that only the coworker network has significant effect. An increase in the average participation of coworkers leads to a decrease in the likelihood of household exiting the stock market. Neighbors and Family networks do not have any significant effect on investment behavior of individual households when we pool all networks in one estimation.

5.1 Networks and Time

In the previous analysis we assume that endogenous effects do not vary over time. This corresponds to assuming that over the business cycle households pay the same attention to coworkers, family, and neighbors investment decisions. If we relax that assumption, we can observe time fluctuation in importance of each network. To test this hypothesis, we estimate several dynamic models. The first model introduces interactions of the average stock market participation in networks with year dummies. This allows us to see if the importance of various networks changes over time. The estimation equation is:

$$
\begin{aligned}
\triangle b_{i,t} = {} & \alpha_0 \\
& + \alpha_1 \triangle \overline{b_{i,t}^{CW}} + \beta_1 \overline{X_{i,t}^{CW}} + \mu_1 \triangle \overline{b_{i,t}^{CW}} YearFE_t \\
& + \alpha_2 \triangle \overline{b_{i,t}^{Neigh}} + \beta_2 \overline{X_{i,t}^{Neigh}} + \mu_2 \triangle \overline{b_{i,t}^{Neigh}} YearFE_t \\
& + \alpha_3 \triangle \overline{b_{i,t}^{Fam}} + \beta_3 \overline{X_{i,t}^{Fam}} + \mu_3 \triangle \overline{b_{i,t}^{Fam}} YearFE_t \\
& + \delta_1 X_{i,t}^{Wife} + \delta_2 X_{i,t}^{Husb} + \gamma_1 YearFE_t + \gamma_2 RegionFE_i + \epsilon_{i,t},
\end{aligned}
\tag{18}
$$

Table 11 presents the result for the OLS and MLOGIT estimation. As we can see, coworkers network is the only type of the network affecting the household stock market participation consistently over time. The average stock market participation of coworkers has positive effect on the likelihood of the household investing in stocks. When the ratio of coworkers investing in stocks increases the probability of the household entering the stock market and decreases the probability of the household exiting. The results indicate that households are following their coworkers when making investment decisions. At the same time, the other types of networks do not have any consistent effect over time. Family networks are slightly significant in 2006 and Neighbors are significant in 2011, but the results are not consistent over time.

In summary, coworkers network has a time-invariant and significant effect on household's stock market participation, while Family and Neighbor networks do not.

6 Conclusion

In this paper we investigate the effect of household's social network on its stock market participation. We look into three types of peers: coworkers, neighbors, and family. We start with analyzing effects of networks separately. The results indicate that all networks have significant effect on household's stock market participation – coworkers and neighbors have strong effect on the probability of household investing in stocks. The results are robust to an instrumental variable strategy that attempts to control for unobservable shocks hitting both the household and its networks. Using instrumented multinomial Logit approach, we also see the effect of each network on the likelihood of household entering (starting investing in) the stock market and on it exiting (closing the account) the stock market. Considered separately, all networks have strong positive effect on the likelihood of entrance, and leave the probability of exit unaffected. These results are consistent with the literature on peer effects on the stock market participation.

However, when we include the effect of all networks simultaneously, only the effect for co-workers remain. There is a strong positive effect of coworkers' beliefs on the probability of entering, and a strong negative effect for the probability of exiting, but the results for neighbors and family networks are insignificant. This indicates that it is important to control for all peer groups simultaneously.

The results of our paper confirm that households do not listen to all their peer equally. Over time households consistently follow stock market decisions of their coworkers and ignore decisions of family members and neighbors.

References

Abel, Andrew B (1990). "Asset prices under habit formation and catching up with the Joneses". In: *The American Economic Review* 80.2, pp. 38–42.

Ang, Andrew, Geert Bekaert, and Jun Liu (2005). "Why stocks may disappoint". In: *Journal of Financial Economics* 76.3, pp. 471–508.

Bikhchandani, Sushil and Sunil Sharma (2000). "Herd behavior in financial markets". In: *IMF Staff papers*, pp. 279–310.

Blume L.E., Brock W.A. Durlauf S.N. and Y.M. Ioannides (2011). "Identification of social interactions". In: *Handbook of social economics*, pp. 853–964.

Bramoullé Y., Djebbari H. and B. Fortin (2009). "Identification of peer effects through social networks". In: *Journal of Econometrics* 150, pp. 41–55.

Brown, Jeffrey R et al. (2008). "Neighbors matter: Causal community effects and stock market participation". In: *The Journal of Finance* 63.3, pp. 1509–1531.

Brown, Sarah and Karl Taylor (2010). "Social interaction and stock market participation: evidence from British Panel Data". In:

Bursztyn, Leonardo et al. (2014). "Understanding mechanisms underlying peer effects: Evidence from a field experiment on financial decisions". In: *Econometrica* 82.4, pp. 1273–1301.

Calvet, Laurent E, John Y Campbell, and Paolo Sodini (2009). *Measuring the financial sophistication of households*. Tech. rep. National Bureau of Economic Research.

Calvó-Armengol, Antoni, Eleonora Patacchini, and Yves Zenou (2009). "Peer effects and social networks in education". In: *The Review of Economic Studies* 76.4, pp. 1239–1267.

Campbell, John Y (2006). "Household finance". In: *The Journal of Finance* 61.4, pp. 1553–1604.

Campbell, John Y and John H Cochrane (1999). “Vby force of habit: A consumption based explanation of aggregate stock market behaviorv”. In: *Journal of Political Economy* 107.2, pp. 205–251.

Cao, H Henry, Tan Wang, and Harold H Zhang (2005). “Model uncertainty, limited market participation, and asset prices”. In: *Review of Financial Studies* 18.4, pp. 1219–1251.

Changwony, Frederick K, Kevin Campbell, and Isaac T Tabner (2014). “Social engagement and stock market participation*”. In: *Review of Finance*, rft059.

Chari, Varadarajan V and Patrick J Kehoe (2004). “Financial crises as herds: overturning the critiques”. In: *Journal of Economic Theory* 119.1, pp. 128–150.

Christiansen, Charlotte, Juanna Schröter Joensen, and Jesper Rangvid (2008). “Are economists more likely to hold stocks?” In: *Review of Finance* 12.3, pp. 465–496.

De Giorgi, Giacomo, Anders Frederiksen, and Luigi Pistaferri (2015). *Consumption network effects*.

De Giorgi, Giacomo, Michele Pellizzari, and Silvia Redaelli (2010). “Identification of social interactions through partially overlapping peer groups”. In: *American Economic Journal: Applied Economics*, pp. 241–275.

DeMarzo, Peter M, Ron Kaniel, and Ilan Kremer (2004). “Diversification as a public good: Community effects in portfolio choice”. In: *The Journal of Finance* 59.4, pp. 1677–1716.

— (2008). “Relative wealth concerns and financial bubbles”. In: *Review of Financial Studies* 21.1, pp. 19–50.

Dow, James and Sergio Ribeiro da Costa Werlang (1992). “Uncertainty aversion, risk aversion, and the optimal choice of portfolio”. In: *Econometrica: Journal of the Econometric Society*, pp. 197–204.

Duflo, Esther and Emmanuel Saez (2002). “Participation and investment decisions in a retirement plan: The influence of colleagues choices”. In: *Journal of public Economics* 85.1, pp. 121–148.

Durlauf, S.N. and Y.M. Ioannides (2010). "Social interactions". In: *Annual Review of Economics* 2, pp. 451–478.

Epstein, Larry G and Martin Schneider (2007). "Learning under ambiguity". In: *The Review of Economic Studies* 74.4, pp. 1275–1303.

Gali, Jordi (1994). "Keeping up with the Joneses: Consumption externalities, portfolio choice, and asset prices". In: *Journal of Money, Credit and Banking* 26.1, pp. 1–8.

Guiso, Luigi and Tullio Jappelli (2005). "Awareness and stock market participation". In: *Review of Finance* 9.4, pp. 537–567.

Guiso, Luigi, Paola Sapienza, and Luigi Zingales (2008). "Trusting the stock market". In: *the Journal of Finance* 63.6, pp. 2557–2600.

Heaton, John and Deborah Lucas (2000). "Portfolio choice and asset prices: The importance of entrepreneurial risk". In: *The journal of finance* 55.3, pp. 1163–1198.

Hong, Harrison, Jeffrey D Kubik, and Jeremy C Stein (2004). "Social interaction and stock-market participation". In: *The journal of finance* 59.1, pp. 137–163.

Kaustia, Markku and Samuli Knüpfer (2012). "Peer performance and stock market entry". In: *Journal of Financial Economics* 104.2, pp. 321–338.

Kelejian, Harry H and Ingmar R Prucha (2010). "Specification and estimation of spatial autoregressive models with autoregressive and heteroskedastic disturbances". In: *Journal of Econometrics* 157.1, pp. 53–67.

Li, Geng (2014). "Information sharing and stock market participation: Evidence from extended families". In: *Review of Economics and Statistics* 96.1, pp. 151–160.

Lin, Xu (2010). "Identifying peer effects in student academic achievement by spatial autoregressive models with group unobservables". In: *Journal of Labor Economics* 28.4, pp. 825–860.

Liu, Xiaodong and Lung-fei Lee (2010). "GMM estimation of social interaction models with centrality". In: *Journal of Econometrics* 159.1, pp. 99–115.

Mankiw, N Gregory and Stephen P Zeldes (1991). "The consumption of stockholders and nonstockholders". In: *Journal of financial Economics* 29.1, pp. 97–112.

Manski, Charles F (1993). "Identification of endogenous social effects: The reflection problem". In: *The review of economic studies* 60.3, pp. 531–542.

Mehra, Rajnish and Edward C Prescott (1985). "The equity premium: A puzzle". In: *Journal of monetary Economics* 15.2, pp. 145–161.

Merton, Robert C (1969). "Lifetime portfolio selection under uncertainty: The continuous-time case". In: *The review of Economics and Statistics*, pp. 247–257.

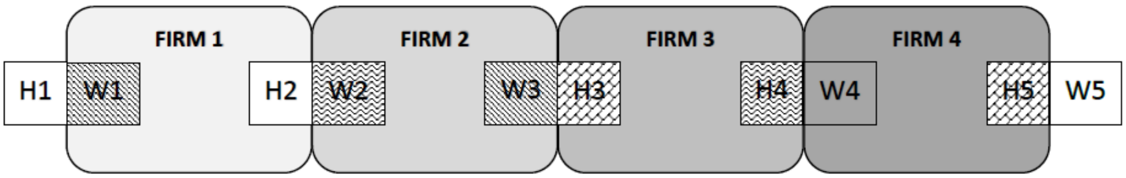
Patacchini, Eleonora and Edoardo Rainone (2014). "The word on banking-social ties, trust, and the adoption of financial products". In: *Tech. Rep., Einaudi Institute for Economics and Finance (EIEF).*

Van Rooij, Maarten, Annamaria Lusardi, and Rob Alessie (2011). "Financial literacy and stock market participation". In: *Journal of Financial Economics* 101.2, pp. 449–472.

Vissing-Jorgensen, Annette (2002). *Limited asset market participation and the elasticity of intertemporal substitution*. Tech. rep. National Bureau of Economic Research.

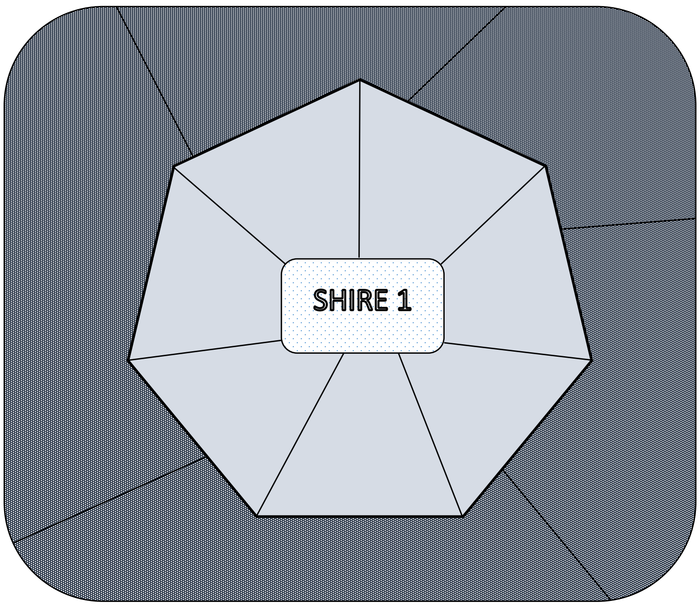
Figures

Figure 1: Example of identification of Instruments for "colleagues"



In Figure 1 peers that are instruments for each other are matched by the same pattern. All firms are different firms and all households are married couples, where both spouses are employed in different firms

Figure 2: Example of identification of Instruments for "neighbours"



In Figure 2 peers that would be instruments for the closest neighbours are matched by pattern. The first distance neighbours live in the same shire, the second distance neighbours live in the second closest shires (matched by light gray), and the third distance neighbours are located in the third furthest shires (marked by dark color).

Figure 3: Shires in Denmark

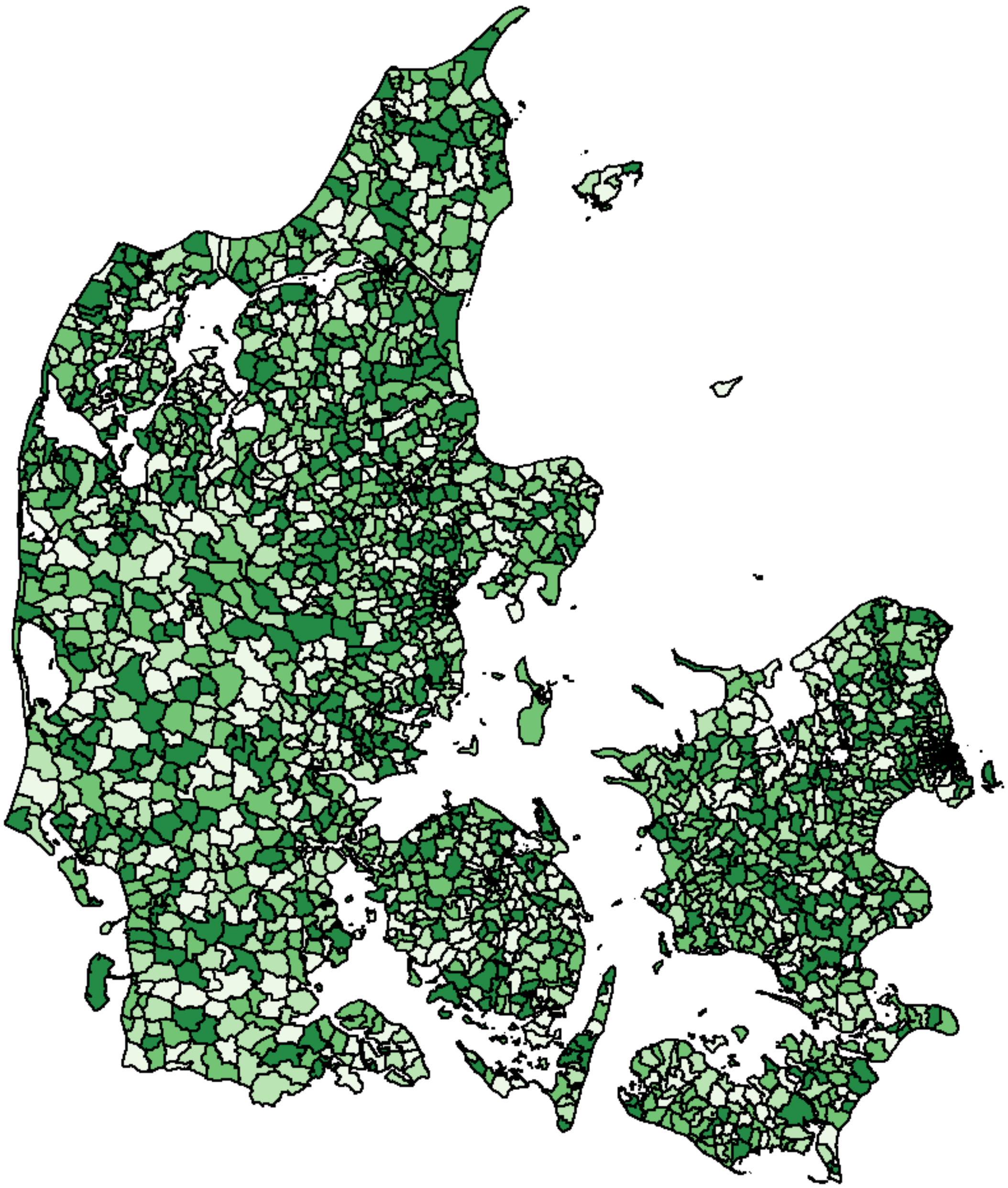


Figure 4: Example of identification of Close Family and Family Instruments

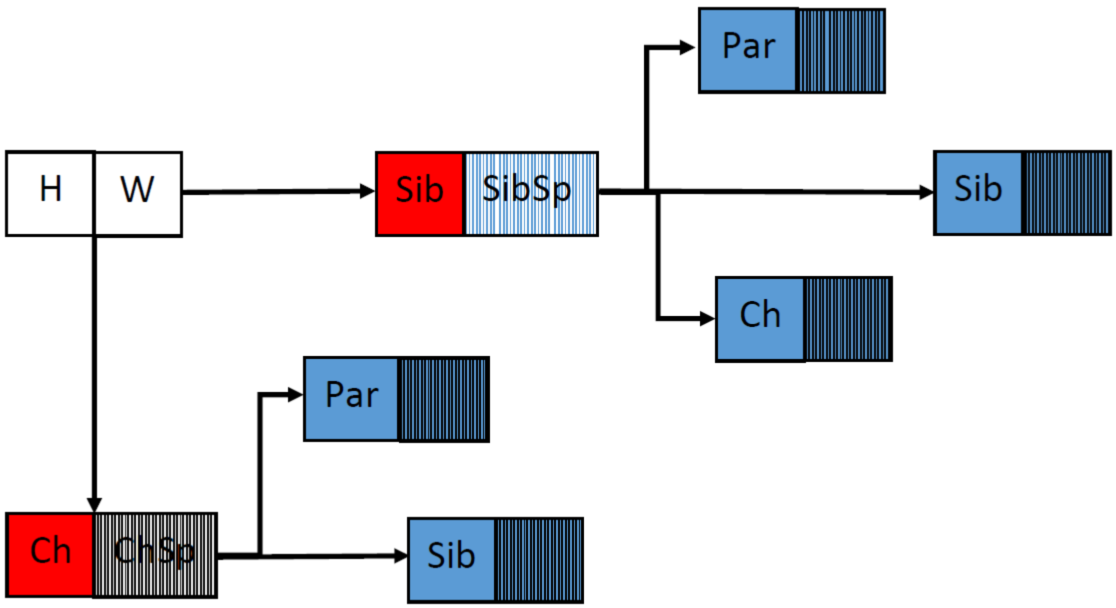


Figure 4 represents a "family tree". The close family is highlighted in red. Instruments for family are highlighted with blue and blue with pattern.

Tables

Table 1: Descriptive Statistics of Individuals, Colleagues, and Coworker Instruments

	Individuals	Colleagues	Coworker Instruments
Change in SMP	0.00	0.00	
	(0.18)	(0.17)	
Log Financial Wealth	10.16	10.14	10.25
	(2.60)	(1.22)	(0.61)
Log Labor Income	12.72	12.71	12.75
	(0.54)	(0.32)	(0.16)
Age	45.13	44.41	44.03
	(8.94)	(4.26)	(1.43)
Years of Schooling	13.16	13.20	13.74
	(3.10)	(1.88)	(0.94)
Share of Females	0.50	0.51	0.51
	(0.50)	(0.32)	(0.13)
Job Tenure	6.11	5.74	6.05
	(6.68)	(3.79)	(1.53)
Occupation: Blue Collar	0.13	0.13	0.11
	(0.34)	(0.25)	(0.12)
Occupation: White Collar	0.63	0.64	0.73
	(0.48)	(0.34)	(0.15)
Occupation: Manager	0.05	0.04	0.03
	(0.21)	(0.09)	(0.03)
# of Kids 0-6	0.39	0.42	0.45
	(0.69)	(0.30)	(0.09)
# of Kids 7-18	0.77	0.75	0.76
	(0.93)	(0.36)	(0.12)
Sector: Manufacturing	0.13	0.14	
	(0.34)	(0.34)	
Sector: Services	0.62	0.63	
	(0.49)	(0.47)	
Sector: Construction	0.04	0.04	
	(0.20)	(0.19)	
Sector: Other	0.21	0.20	
	(0.41)	(0.39)	
Firm Size	403.19	428.19	1,884.29
	(1,034.80)	(1,061.90)	(1,351.76)
Firm Growth	0.38	0.36	0.28
	(1.78)	(1.73)	(0.75)
Firm Type: Public Sector	0.29	0.30	0.34
	(0.45)	(0.46)	(0.26)
Firm Type: Publicly Traded	0.40	0.41	0.35
	(0.49)	(0.49)	(0.23)
Firm Type: Limited Liability	0.07	0.05	0.02
	(0.25)	(0.23)	(0.07)
FirmType: Other	0.25	0.23	0.29
	(0.43)	(0.42)	(0.25)
Observations	5,245,994	5,245,994	4,164,938

Table 2: Descriptive Statistics: Individuals, Neighbors, and Instruments for Neighbors

	Individuals of Interest	Neighbors	Neighbors Instruments
Change in SMP	0.01	0.00	
	(0.19)	(0.05)	
Log Financial Wealth	10.21	10.14	10.13
	(2.60)	(0.47)	(0.46)
Log Labor Income	12.73	12.71	12.70
	(0.55)	(0.12)	(0.11)
Age	45.39	44.61	44.52
	(8.66)	(1.10)	(1.04)
Gender	0.50	0.50	0.50
	(0.50)	(0.02)	(0.02)
Years of Schooling	13.17	13.12	13.11
	(3.08)	(0.62)	(0.57)
Job Tenure	6.35	5.80	5.75
	(6.85)	(0.69)	(0.65)
Occupation: Blue	0.13	0.13	0.14
	(0.34)	(0.05)	(0.05)
Occupation: White	0.62	0.61	0.60
	(0.49)	(0.06)	(0.06)
Occupation: Manager	0.05	0.05	0.04
	(0.22)	(0.02)	(0.01)
# of Kids 0-6	0.38	0.42	0.42
	(0.69)	(0.06)	(0.05)
# of Kids 7-18	0.79	0.77	0.77
	(0.94)	(0.11)	(0.11)
Sector: Manufacturing	0.14	0.14	
	(0.34)	(0.04)	
Sector: Services	0.60	0.60	
	(0.49)	(0.06)	
Sector: Construction	0.04	0.04	
	(0.20)	(0.02)	
Sector: Other	0.22	0.23	
	(0.41)	(0.06)	(0.06)
Firm Size	410.76	400.70	398.64
	(1,048.62)	(132.45)	(120.63)
Firm Growth	0.37	0.44	0.44
	(1.77)	(0.23)	(0.23)
Firm Type: Public Sector	0.28	0.27	0.26
	(0.45)	(0.05)	(0.07)
Firm Type: Publicly Traded	0.40	0.38	0.37
	(0.49)	(0.05)	(0.08)
Firm Type: Limited Liability	0.08	0.08	0.07
	(0.27)	(0.02)	(0.02)
Firm Type: Other	0.24	0.23	0.22
	(0.43)	(0.04)	(0.05)
Observations	6,776,832	6,776,832	6,723,399

Table 3: Descriptive Statistics: Individuals, Family, and Instruments for Family

	Individuals of Interest	Family	Instruments
Change in SMP	0.01	0.01	
	(0.20)	(0.29)	
Log Financial Wealth	10.28	9.88	10.23
	(2.67)	(2.42)	(2.37)
Log Labor Income	12.69	12.70	12.66
	(0.53)	(0.50)	(0.53)
Age	48.87	35.64	47.15
	(9.16)	(6.99)	(9.22)
Gender	0.50	0.49	0.47
	(0.50)	(0.50)	(0.41)
Years of Schooling	12.74	13.67	12.76
	(3.06)	(2.67)	(2.83)
Job Tenure	7.42	4.03	6.38
	(7.57)	(4.81)	(6.33)
Occupation: Blue	0.15	0.14	0.14
	(0.35)	(0.34)	(0.31)
Occupation: White	0.59	0.66	0.59
	(0.49)	(0.47)	(0.44)
Occupation: Manager	0.05	0.04	0.04
	(0.22)	(0.18)	(0.18)
# of Kids 0-6	0.28	0.92	0.34
	(0.62)	(0.85)	(0.59)
# of Kids 7-18	0.54	0.63	0.55
	(0.86)	(0.88)	(0.81)
Sector: Manufacturing	0.14	0.15	
	(0.35)	(0.35)	
Sector: Services	0.59	0.57	
	(0.49)	(0.49)	
Sector: Construction	0.05	0.05	
	(0.21)	(0.22)	
Sector: Other	0.22	0.23	
	(0.41)	(0.42)	
Firm Size	371.69	413.48	367.25
	(985.91)	(1,052.98)	(874.67)
Firm Growth	0.37	0.44	0.44
	(1.76)	(1.91)	(1.69)
Firm Type: Public Sector	0.28	0.25	0.28
	(0.45)	(0.43)	(0.40)
Firm Type: Publicly Traded	0.39	0.42	0.37
	(0.49)	(0.49)	(0.43)
Firm Type: Limited Liability	0.08	0.08	0.08
	(0.28)	(0.28)	(0.24)
Firm Type: Other	0.24	0.22	0.24
	(0.43)	(0.41)	(0.38)
Observations	2,097,776	569,744	569,744

Table 4: Coworkers Network

	(1)	(2)
	OLS	IVREG
Change in Coworkers SMP	0.0431***	0.122***
	(49.46)	(3.56)
Age Coworkers	0.0000249	0.000208
	(0.05)	(0.34)
Age Squared Coworkers	-0.000000436	-0.00000444
	(-0.08)	(-0.63)
Male Ratio among Coworkers	-0.00285***	-0.00267***
	(-4.30)	(-3.45)
Coworkers' Kids younger than 6	0.000568	0.000111
	(0.78)	(0.13)
Coworkers' Kids 7-18	0.000689	0.000959
	(1.47)	0.000959
Length of Education, Coworkers	-0.000315**	-0.000506**
	(-2.80)	(-3.08)
Coworkers' Occupation: Blue Collar	-0.000697	0.000215
	(-0.68)	(0.18)
Coworkers' Occupation: White Collar	0.00120	0.000889
	(1.56)	(0.96)
Coworkers' Occupation: Management	-0.00574**	-0.00705**
	(-3.14)	(-3.14)
Demograhic IVs	No	Yes
Firm IVs	No	Yes
First Stage F-statistic	.	125.4
Observations	2,059,900	2,059,900

Note: *,**,***= significant at 5%, 1%, 0.1%. T-statistics in parenthesis. Dependent variable is the Change in household stock market participation, where -1 indicates that the household exits the the stock market, 0 indicates no change in stock market participation, and 1 indicates that the household entered the stock market. Individual controls (separately for husband and wife) include: Age, Age sq., Years of schooling, Occupation dummies, Industry dummies, Public sector dummy, Firm size, Firm growth, Firm type dummies. Household controls: Region dummies, # kids 0-6, # kids 7-18. Demographic IVs: Age, Age sq:, Years of schooling, # kids0-6, # kids 7-18, share of male peers, shares of peers by occupation, Log of Income. Firm IVs: Public sector dummy, Firm type dummy. We also control for year fixed-effects. For details on the weighting schemes see the main text.

Table 5: Coworkers Network: MLOGIT

	(1)	(2)
	Exit	Entrance
Change in Coworkers SMP (predicted)	0.225	1.951^{***}
	(0.36)	(3.59)
Observations	2,065,994	2,065,994

Note: *,**,***= significant at 5%, 1%, 0.1%. Dependent variable: Change in household stock market participation ((-1) - Exit from the stock market, (0) - No change in stock market participation, (1) - Entrance to the stock market). No change outcome is used s a base outcome. Individual controls (separately for husband and wife): Age, Age sq., Years of schooling, Occupation dummies, Industry dummies, Public sector dummy, Firm size, Firm growth, Firm type dummies. Household controls: Region dummies, # kids 0-6, # kids 7-18. Demographic IVs: Age, Age sq:, Years of schooling, # kids0-6, # kids 7-18, share of male peers, shares of peers by occupation, Log of Income. Firm IVs: Public sector dummy, Firm type dummy. We also control for yearfixed effects. For details on the weighting schemes see the main text.

Table 6: Family Network

	(1)	(2)
	OLS	IVREG
Change in Family SMP	0.00859***	0.433
	(6.12)	(1.35)
Age Family	0.000461	0.00106
	(0.66)	(1.11)
Age Squared Family	-0.00000559	-0.0000178
	(-0.64)	(-1.28)
Male Ration Among Family	-0.000234	-0.000128
	(-0.25)	(0.11)
Length of Education, Family	-0.0000118	-0.000596
	(-0.06)	(-1.20)
Occupation Family: Blue Collar	0.000805	0.00162
	(0.50)	(0.8)
Occupation Family: White Collar	0.000691	-0.00272
	(0.54)	(-0.91)
Occupation Family: Management	-0.00223	-0.00994
	(-0.90)	(-1.52)
Coworkers' Kids younger than 6	0.000339	0.00129
	(0.54)	(1.24)
Coworkers' Kids age of 7-18	0.000496	0.00116
	(0.80)	(1.30)
IV: Length of Education	No	Yes
First Stage F-statistic	.	6.007
Observations	223,181	223,181

Note: *,**,***= significant at 5%, 1%, 0.1%. Dependent variable: Change in household stock market participation ((-1) - Exit from the stock market, (0) - No change in stock market participation, (1) - Entrance to the stock market). Individual controls (separately for husband and wife): Age, Age sq., Years of schooling, Occupation dummies, Industry dummies, Public sector dummy, Firm size, Firm growth, Firm type dummies. Household controls: Region dummies, # kids 0-6, # kids 7-18. Only one IV variable works for the Family network: Length of education of family members. We also control for year fixed effects. For details on the weighting schemes see the main text.

Table 7: Coworkers Network: MLOGIT

	(1) Exit	(2) Entrance
Change in Coworkers SMP (predicted)	-11.20	55.14*
	(-0.38)	(2.03)
Observations	223,181	223,181

Note: *,**,***= significant at 5%, 1%, 0.1%. Dependent variable: Change in household stock market participation ((-1) - Exit from the stock market, (0) - No change in stock market participation, (1) - Entrance to the stock market). No change outcome is used s a base outcome. Individual controls (separately for husband and wife): Age, Age sq., Years of schooling, Occupation dummies, Industry dummies, Public sector dummy, Firm size, Firm growth, Firm type dummies. Household controls: Region dummies, # kids 0-6, # kids 7-18. Exogenous Instrument: Length of Education of family members. We also control for yearfixed effects. For details on the weighting schemes see the main text.

Table 8: Neighbors Network: OLS and IVREG

	(1) OLS	(2) IVREG
Change in Neighbors SMP	0.105***	0.287**
	(40.87)	(2.86)
Age of Neighbors	0.0117**	0.0217**
	(2.87)	(3.18)
Age Square of Neighbors	-0.000140**	-0.000264**
	(-3.00)	(-3.21)
Male Ratio among Neighbors	0.00226	0.00351
	(0.38)	(0.59)
Lenght of Education of Neighbors	-0.000987*	-0.00168**
	(-2.21)	(-2.88)
Occupation Neighbors: Blue Colar	-0.00171	-0.00653
	(-0.29)	(-0.99)
Occupation Neighbors: White Collar	-0.00140	-0.0158
	(-0.30)	(-1.71)
Occupation Neighbors: Management	-0.0221*	-0.0508**
	(-2.08)	(-2.64)
Kids of Neighbors younger than 6	-0.0127***	-0.0169***
	(-4.33)	-4.57
Kids of Neighbors 7-18	0.000103	-0.000557
	(0.08)	(-0.39)
Demograhic IVs	No	Yes
Firm IVs	No	Yes
First Stage F-statistic	.	280.1
Observations	3,381,537	3,381,537

Note: *,**,***= significant at 5%, 1%, 0.1%. Dependent variable: Change in household stock market participation ((-1) - Exit from the stock market, (0) - No change in stock market participation, (1) - Entrance to the stock market). Individual controls (separately for husband and wife): Age, Age sq., Years of schooling, Occupation dummies, Industry dummies, Public sector dummy, Firm size, Firm growth, Firm type dummies. Household controls: Region dummies, # kids 0-6, # kids 7-18. Demographic IVs: Age, Age sq:, Ration of Neighbors occupied in Management # kids 7-18, Log of Income. Firm IVs: Firm Size, LTD Firms Ratio, Publicly Tarded Firm Ratio. We also control for year fixed effects. For details on the weighting schemes see the main text.

Table 9: Neighbors Network: MLOGIT

	(1)	(2)
	Exit	Entrance
Change in Coworkers SMP (predicted)	-2.280	12.77**
	(-0.51)	(3.21)
Observations	3,381,537	3,381,537

Note: *,**,***= significant at 5%, 1%, 0.1%. Dependent variable: Change in household stock market participation ((-1) - Exit from the stock market, (0) - No change in stock market participation, (1) - Entrance to the stock market). No change outcome is used s a base outcome. Individual controls (separately for husband and wife): Age, Age sq., Years of schooling, Occupation dummies, Industry dummies, Public sector dummy, Firm size, Firm growth, Firm type dummies. Household controls: Region dummies, # kids 0-6, # kids 7-18. Exogenous Demographic IVs: Age, Age sq:, Ration of Neighbors occupied in Management # kids 7-18, Log of Income. Firm IVs: Firm Size, LTD Firms Ratio, Publicly Tarded Firm Ratio. We also control for yearfixed effects. For details on the weighting schemes see the main text.

Table 10: All Networks: OLS and IVREG

	(1)	(2)	(3)
	OLS	MLOGIT:Entrance	MLOGIT:Exit
Change in Family SMP	0.00792***	-2.927	-15.29
	(4.24)	(-0.26)	(-1.17)
Change in Coworkers SMP	0.0410***	1.309	-9.287*
	(9.78)	(0.28)	(-1.73)
Change in Neighbors SMP	0.128***	11.23	12.01
	(8.88)	(0.65)	(0.60)
Demograhic IVs	No	Yes	Yes
Firm IVs	No	Yes	Yes
Observations	146,833	146,833	146,833

Note: *,**,***= significant at 10%, 5%, 1%. Dependent variable: Change in household stock market participation ((-1) - Exit from the stock market, (0) - No change in stock market participation, (1) - Entrance to the stock market). Individual controls (separately for husband and wife): Age, Age sq., Years of schooling, Occupation dummies, Industry dummies, Public sector dummy, Firm size, Firm growth, Firm type dummies. Household controls: Region dummies, # kids 0-6, # kids 7-18. We use separatle IVs for Coworkers, Neighbors, and Family defined before. We also control for year fixed effects. For details on the weighting schemes see the main text.

Table 11: All Networks: OLS and IVREG

	(1) OLS	(2) MLOGIT:Exit	(3) MLOGIT:Entrance
Change in Family SMP	0.00318	-15.73	0.401
	(0.50)	(-1.11)	(0.03)
Change in Coworkers SMP	0.0419**	-19.63**	17.72**
	(3.01)	(-2.59)	(2.60)
Change in Neighbors SMP	0.09	2.228	3.6
	(1.79)	(0.11)	(0.19)
2002 ⋂ Change in Family SMP	0.000282	7.245	-5.157
	(0.03)	(0.64)	(-0.57)
2003 ⋂ Change in Family SMP	-0.00392	-0.541	-0.227
	(-0.43)	(-0.05)	(-0.02)
2004 ⋂ Change in Family SMP	0.00432	-7.365	-7.83
	(0.49)	(-0.73)	(-0.77)
2005 ⋂ Change in Family SMP	0.00204	-1.216	-5.887
	(0.24)	(-0.13)	(-0.70)
2006 ⋂ Change in Family SMP	0.0201*	2.907	-0.156
	(2.45)	(0.30)	(-0.02)
2007 ⋂ Change in Family SMP	0.00714	11.75	-8.974
	(0.82)	(1.23)	(-1.03)
2008 ⋂ Change in Family SMP	0.00624	0.412	0.484
	(0.65)	(0.04)	(0.05)
2009 ⋂ Change in Family SMP	0.00757	6.346	4.599
	(0.70)	(0.56)	(0.39)
2010 ⋂ Change in Family SMP	0.00825	8.694	-2.893
	(0.76)	(0.76)	(-0.27)
2011 ⋂ Change in Family SMP	0.0062	-11.56	-4.98
	(0.60)	(-1.08)	(-0.43)
2012 ⋂ Change in Family SMP	-0.00914	-7.101	-0.039
	(-0.88)	(-0.67)	(-0.00)
2013 ⋂ Change in Family SMP	0.00297	-2.198	0.143
	(0.27)	(-0.20)	(0.01)
2002 ⋂ Change in Coworkers SMP	0.00791	23.43*	-12.97
	(0.41)	(2.54)	(-1.69)
2003 ⋂ Change in Coworkers SMP	-0.00952	4.479	-24.87**
	(-0.48)	(0.48)	(-2.70)
2004 ⋂ Change in Coworkers SMP	0.00489	2.275	-24.51**
	(0.25)	(0.24)	(-2.74)
2005 ⋂ Change in Coworkers SMP	0.0642***	18.34*	-3.277
	(3.45)	(2.15)	(-0.46)
2006 ⋂ Change in Coworkers SMP	-0.013	15.04	-9.38
	(-0.70)	(1.73)	(-1.35)
2007 ⋂ Change in Coworkers SMP	-0.0099	1.194	-25.84***
	(-0.53)	(0.13)	(-3.36)
2008 ⋂ Change in Coworkers SMP	-0.0161	6.508	-16.9
	(-0.77)	(0.68)	(-1.72)
2009 ⋂ Change in Coworkers SMP	-0.0124	8.359	-13.21
	(-0.51)	(0.72)	(-1.15)
2010 ⋂ Change in Coworkers SMP	-0.0367	18.03	-22.88*
	(-1.51)	(1.48)	(-2.38)
2011 ⋂ Change in Coworkers SMP	0.00648	28.96*	-30.02**
	(0.27)	(2.57)	(-2.73)
2012 ⋂ Change in Coworkers SMP	-0.0436	14.19	-43.68***
	(-1.76)	(1.21)	(-4.40)
2013 ⋂ Change in Coworkers SMP	-0.0329	1.658	-30.59
	(-1.28)	(0.14)	(-1.92)
2002 ⋂ Change in Neighbors SMP	0.0388	4.033	4.755
	(0.58)	(0.26)	(0.43)
2003 ⋂ Change in Neighbors SMP	-0.00825	25.09	26.31*
	(-0.12)	(1.96)	(2.02)
2004 ⋂ Change in Neighbors SMP	-0.0249	14.95	16.23
	(-0.38)	(1.14)	(1.32)
2005 ⋂ Change in Neighbors SMP	0.0251	4.067	-4.894
	(0.39)	(0.34)	(-0.46)
2006 ⋂ Change in Neighbors SMP	0.0783	9.4	7.106
	(1.22)	(0.70)	(0.65)
2007 ⋂ Change in Neighbors SMP	0.0419	25.79	16.05
	(0.67)	(1.92)	(1.38)
2008 ⋂ Change in Neighbors SMP	0.0446	11.73	10.51
	(0.56)	(0.77)	(0.69)
2009 ⋂ Change in Neighbors SMP	-0.0227	6.521	17.03
	(-0.26)	(0.39)	(0.95)
2010 ⋂ Change in Neighbors SMP	0.109	3.481	5.52
	(1.23)	(0.20)	(0.43)
2011 ⋂ Change in Neighbors SMP	0.183*	11.52	10.13
	(2.14)	(0.74)	(0.66)
2012 ⋂ Change in Neighbors SMP	0.154	8.717	36.21*
	(1.74)	(0.56)	(2.54)
2013 ⋂ Change in Neighbors SMP	0.13	19.73	-6.239
	(1.44)	(1.19)	(-0.29)
Demograhic IVs	No	Yes	Yes
Firm IVs	No	Yes	Yes
Year FE	Yes	Yes	Yes
Region FE	Yes	Yes	Yes
Observations	146,833	146,833	146,833

Note: *,**,***= significant at 5%, 1%, 0.1%. We use separatle IVs for Coworkers, Neighbors, and Family defined before.

Appendix 1: Identification of networks

1. Identifying Co-workers social network: G

We consider peers of the 3^{rd} degree: *co-workers of the spouse of individual i's co-workers.*

1.1 Co-workers Network: CW

1. Create dummy variables for each workplace that we have in the sample. Let's assume that number of different workplaces is W. As a result we have $Work\_dummy_m \ (2N \times 1)$, where $m = 1...W$. and N is a total number of households in our sample.

2. For each m (workplace identificator) we create a co-workers network: CW_m, where

$$\underset{(2N\times 2N)}{CW_m} = \underset{(2N\times 1)}{[Work\_dummy_m]}\underset{(1\times 2N)}{[Work\_dummy_m]'}$$

$$CW_{m,ij} = \begin{cases} 1 & \text{if } i \text{ and } j \text{ are co-workers at the firm } m \\ 0 & \text{o/w} \end{cases}$$

3. To construct the final matrix of co-workers we need to sum up matrices for each firm m, i.e.

$$CW = \sum_m CW_m$$

As a result we have a co-workers network matrix $CW \ (2N \times 2N)$, such that

$$CW_{ij} = \begin{cases} 1 & \text{if } i \text{ and } j \text{ are co-workers} \\ 0 & \text{o/w} \end{cases}$$

1.2 Spousal network

Using gender and household identifiers in the data we can determine a wife and a husband for the same household. Using identifiers of wives and husbands, we can construct matrix $SP \ (2N \times 2N)$, such that

$$SP_{ij} = \begin{cases} 1 & \text{if } i \text{ and } j \text{ are spouses} \\ 0 & \text{o/w} \end{cases}$$

1.3 Social Network: G

For our purposes we are looking at peers of 3^{rd} degree. In that case the social network can be constructed in the following way:

$$\widetilde{G} = CW \times SP \times CW \tag{19}$$

where

$$\widetilde{G}_{ij} = \begin{cases} 1 & \text{if } i \text{ is a third degree peer for } j \text{ from co-workers network} \\ 0 & \text{o/w} \end{cases}$$

The Equation 19 above describes instruments for co-workers (peers in the co-worker network). One of the examples of the instrument would be a co-worker of a spouse of a co-worker.

Final result should be a normalized social network, G, such that

$$G_{ij} = \begin{cases} \frac{\widetilde{G}_{ij}}{\sum_j^{2N} \widetilde{G}_{ij}} & \text{if } \widetilde{G}_{ij} \neq 0 \\ 0 & \text{o/w} \end{cases}$$

2. Identifying wife and husband's induced co-workers networks: G^w and G^h

2.1 Identify husbands and wives for corresponding households

- S^h $(2N \times N)$ matches husbands with households they are from:

$$S^h_{ij} = \begin{cases} 1 & \text{if } i \text{ is a husband in a household } j \\ 0 & \text{o/w} \end{cases}$$

- S^w $(2N \times N)$ matches wives with households they are from:

$$S^w_{ij} = \begin{cases} 1 & \text{if } i \text{ is a wife in a household } j \\ 0 & \text{o/w} \end{cases}$$

2.2 Wife/Husband-induced- co-workers networks

- G^h $(N \times N)$ husband induced household network: $G^h = (S'_h + S'_w)\, G' S_h$
- G^w $(N \times N)$ wife induced household network: $G^w = (S'_h + S'_w)\, G' S_w$

Family Network

Looking at that kind of network we consider the following relatives as close family: parents, children, and siblings.

1. Close family

1. Matrix of siblings, SIB $(2N \times 2N)$:

$$SIB_{ij} = \begin{cases} 1 & \text{if } i \text{ is a sibling of } j \\ 0 & \text{o/w} \end{cases}$$

2. Matrix of children, $CHILD$ $(2N \times 2N) = PAR'$:

$$CHILD_{ij} = \begin{cases} 1 & \text{if } i \text{ is a child of of } j \\ 0 & \text{o/w} \end{cases}$$

3. Close Family network, $\widetilde{G}$:

$$\widetilde{G_F} = SIB + CHILD$$

4. Normalized Close Family Network, G :

$$G_{F,ij} = \begin{cases} \frac{\widetilde{G}_{F,ij}}{\sum_j^{2N} \widetilde{G}_{F,ij}} & \text{if } G_{F,ij} \neq 0 \\ 0 & \text{o/w} \end{cases}$$

2. Family Instruments

To identify instruments for peers in the family network we need to idetify in-laws (spouses of siblings) and spouses of kids. For those individuals then we identify siblings, aprents and children. Latter will define third-distance family members. The following types of relatives would be a third-distance relatives, for example:

1. Parents of In-Laws, PAR_{INLAWS}

$$PAR_{INLAWS} = INLAWS * PAR$$

2. Siblings of In-Laws, SIB_{INLAWS}

$$SIB_{INLAWS} = SIB * INLAWS$$

3. Children's In-Laws, CH_{INLAWS}

$$CH_{INLAWS} = CH * INLAWS$$

4. Parents of spouses of children, PAR_{SPCH}

$$PAR_{SPCH} = CHILD * SP * PAR$$

5. Children of Children, CH_{CH}

$$CH_{CH} = CH * CH$$

6. Instrumented Family network, $\widetilde{G}$:

$$\widetilde{G_{FamI}} = PAR_{INLAWS} + SIB_{INLAWS} + CH_{INLAWS} + PAR_{SPCH} + CH_{SPCH}$$

7. Normalized Close Family Network, G :

$$G_{FI,ij} = \begin{cases} \frac{\widetilde{G}_{FI,ij}}{\sum_j^{2N} \widetilde{G}_{FI,ij}} & \text{if } G_{FI,ij} \neq 0 \\ 0 & \text{o/w} \end{cases}$$

2.1 Wife/Husband-induced- family networks

- G^h $(N \times N)$ husband induced household network: $G^h_{FI} = (S'_h + S'_w)\, G' S_h$
- G^w $(N \times N)$ wife induced household network: $G^w_{FI} = (S'_h + S'_w)\, G' S_w$

3. Identifying Neighbors social network: G

We consider peers of individual as *neighbors from the same shire*

3.1 Neighbors Network: G_{Neigh}

1. Create dummy variables for each shire that we have in the sample. Let's assume that number of different shires is S. As a result we have $Shire\_dummy_s$ $(2N \times 1)$, where $s = 1...S$. and N is a total number of households in our sample.

2. For each s (shire identificator) we create a neighbors network: $Neigh_s$, where

$$\begin{aligned} \underset{(2N\times 2N)}{Neigh_s} &= \underset{(2N\times 1)}{[Shire\_dummy_s]}\underset{(1\times 2N)}{[Shire\_dummy_s]'} \\ Neigh_{s,ij} &= \begin{cases} 1 & \text{if } i \text{ and } j \text{ are neighbors in the same shire } s \\ 0 & \text{o/w} \end{cases} \end{aligned}$$

3. To construct the final matrix of neighbors we need to sum up matrices for each shire s, i.e.

$$\widetilde{G}_{Neigh} = \sum_{s} Neigh_s$$

As a result we have a co-workers network matrix $\widetilde{G}_{Neigh}$ $(2N \times 2N)$, such that

$$\widetilde{G}_{Neigh,ij} = \begin{cases} 1 & \text{if } i \text{ and } j \text{ are neighbors} \\ 0 & \text{o/w} \end{cases}$$

4. Normalized Neighbors Network, G :

$$G_{Neigh,ij} = \begin{cases} \frac{\widetilde{G}_{Neigh,ij}}{\sum_j^{2N} \widetilde{G}_{Neigh,ij}} & \text{if } \widetilde{G}_{Neigh,ij} \neq 0 \\ 0 & \text{o/w} \end{cases}$$