# How Do Investment Ideas Spread through Social Interaction? Evidence from a Ponzi Scheme[*]

Ville Rantala
University of Miami

December 21, 2016

**Abstract**

A unique dataset from a large Ponzi scheme shows that investment ideas can spread epidemically through social contagion. Investors could join the scheme only by personal invitation from an existing member, and I can observe how the idea spreads from one person to the next based on the inviter-invitee relationships. The social network has so-called scale-free connectivity structure where the distribution of the number of invited people approximately follows a power law. The structure differs significantly from randomly formed networks and explains why word-of-mouth information can spread rapidly even if the average investor does not share it with many others.

*JEL Classification:* D19, G19

[*] E-mail: vrantala@bus.miami.edu. I am grateful to the Finnish National Bureau of Investigations for providing the source documents. I thank Nick Barberis, Peter Bossaerts, James Choi, Nicholas Christakis, Henrik Cronqvist, Florian Ederer, Ester Faia, David Hirshleifer, Harrison Hong, Hans Hvide, Markku Kaustia, Matti Keloharju, Samuli Knüpfer, Alok Kumar, Charlotte Østergaard, Raghu Rau, Sophie Shive, Timo Korkeamäki, Scott Weisbenner, and the participants at the Helsinki Finance Summit 2014, European Finance Association Doctoral Tutorial 2015, European Conference on Household Finance 2015, Conference on Behavioral Aspects of Macroeconomics and Finance 2015, Miami Behavioral Finance Conference 2015, SFS Cavalcade 2016, as well as seminar participants at BI Norwegian Business School, Cass Business School, Copenhagen Business School, Georgetown University, HEC Paris, Indiana University, University of Melbourne, University of Miami, University of New South Wales, and Yale University for helpful comments. I also thank Antti Lehtinen for assistance with geographic information analysis. I acknowledge financial support from the OP Group Research Foundation.

## I. Introduction

Shiller (2000, 2014) proposes that investment ideas can spread like epidemics among investors, and asset prices are influenced by the social dynamics. Do large-scale social contagion effects among investors exist and what does the spreading process of information look like?

In this paper, I study a dataset in which I can directly observe the spreading of an investment idea from one person to the next. The dataset consists of participants of a large investment Ponzi scheme, which investors could join only by personal invitation from an existing member who was referred to as sponsor. This feature allows me to study the spreading and effect of word-of-mouth information at the level of individual people. Information about the scheme was not publicly available, so when a new investor joins, I know that he has learned about the opportunity from the inviter. I can observe the social network of the inviter-invitee relationships and this offers a unique opportunity to take a peek at the process through which an investment idea spreads in the population.

The investment scheme is Wincapita, a Finnish investment operation which was active from 2003 to 2008. Wincapita offered its investors large returns, initially claiming that the profits were generated by sports betting and later by currency trading. In reality, it was a classic Ponzi scheme in which all incoming cash flows came from new and existing investors, and none of the profits were generated by actual trading or investments. The scheme grew very large: In the end, Wincapita had over 10,000 members, which represents approximately 0.2% of the total population of Finland.

The collapse of Wincapita in 2008 lead to one of the largest criminal investigations in Finnish history (YLE News 2008). The dataset of this study has been collected from the investigation documents of the Finnish National Bureau of Investigations. The documents allow me to identify over 5,000 Wincapita investors and I have detailed information of over 3,000 investors who were questioned by the police. In addition to details about the individuals' investments and withdrawals, I also have information about their characteristics, such as age, income, location, and education. I can combine this data with the information on their sponsoring relationships.[1] Because the data was collected from the investors in a formal police interview, it does not suffer from many of the typical reporting and selection biases that can exist in survey data on social relationships. The interviewing officer's responsibility was to collect facts that could be used as evidence in a court proceeding.

Several papers have built theoretical models of the diffusion of information and the structure of information networks among investors,[2] and I can provide empirical evidence on these phenomena. In particular, the dataset allows me to study the following previously unanswered empirical question: What is the social network structure of information diffusion among investors? Empirical social networks connecting people through different kinds of personal ties exhibit strong structural regularities and differ significantly from random graph networks where the connections between nodes are evenly distributed (see Jackson and Rogers 2007 for a review). But because the spreading of word-of-mouth information is typically unobservable, there is still little evidence on whether these commonly observed structures play any role in the diffusion of information within the networks. In social networks, the distribution of connections

[1] There was a financial incentive to sponsor others. Sponsors received 200 euros of (virtual) money for each sponsored investors and 20% of the virtual profits earned by the sponsored investors' investments (the details will be discussed in Section II). Sponsor's sponsor would not receive any portion of the profits, so Wincapita was not a traditional pyramid scheme in which the profits are determined by the investor's level in the "pyramid".

[2] Stein (2008), Han and Yang (2013), and Andrei and Cujean (2015) model the transmission of information through personal communication between investors. Ozsoylev and Walden (2011) study the asset pricing implications of large information networks. Shiller and Pound (1989), Shiller (2000), and Shive (2010) propose that epidemic models can be used to characterize the diffusion of social interest among investors. For general models of word-of-mouth communication see, e.g. Ellison and Fudenberg (1995), Banerjee and Fudenberg (2004), and Cao, Han, and Hirshleifer (2011).

per node typically has a heavy right tail, and the average node to node distances are short. I find that both these characteristics exist in the spreading process of Wincapita.

The distribution of the number of connections per node in Wincapita's sponsoring network has a heavy right tail and it approximately decays as a power law. The empirical probability of sponsoring $k$ investors is proportional to the power of $k$ so that $P(k) \sim k^{-\gamma}$ with a constant $\gamma$. The power-law characteristic is apparent visually in a log-log plot and Kolmogorov-Smirnov test statistics based on a fitted power-law model strongly support the hypothesis that the data follows a power law. Networks with this structure are known as scale-free networks and power-law distributions are very common in empirical social networks (Barabási 2009). Here the power law indicates that a small minority of the investors creates the majority of the social effect.

I compare Wincapita's actual sponsoring network to a simulated random network and find that the power-law topology has a dramatic impact on the spreading rate of information. The random network has the same number of investors, the same percentage of sponsors, and the same average number of sponsored people per sponsor. The only difference compared to the actual network is that the distribution of the number of sponsored people follows a Poisson distribution instead of the power law, as in Erdős and Renyi (1959) model. The comparison shows that an epidemic spreading through the Wincapita network one step at a time reaches all investors in 15 steps. In the simulated network, it takes on average 161 network steps to reach the same number of investors. I also calibrate a simple Ponzi scheme model to both networks and find that the actual network can sustain a significantly higher payout ratio to investors.

I then analyze the spreading dynamics of Wincapita based on network distance. I find that the cumulative number of investors as a function of network distance from the starter of the scheme follows an S-shaped curve. When the scheme grows, the average social distance to the starter of the scheme does not grow at the same rate. Average node-to-node distances in the network are short, even though most people are directly connected to only one or two other people. The S-curve implies that information diffusion within social networks progresses in a nonlinear fashion.

Finally, I study how the characteristics of the inviter are related to the observed sponsoring relationships and the spreading of the scheme. The sponsors have on average higher income than the investor they sponsor (median difference is 9,450 euros per year) and are slightly older (median difference is 1.4 years), suggesting that personal characteristics matter in information diffusion. Most investors join Wincapita after their inviter had already personally generated profits from the scheme. Investors could withdraw funds after being a member for six months, and the median difference in time of joining is seven months. This suggests that peers' personal outcomes are related to the social spreading of investment ideas.

This paper contributes to the literature in several ways. First, I show that a contagious investment idea can spread epidemically in the population through social interaction, gradually affecting larger and larger groups of people, as predicted by Shiller (2000). The existence of behavior contagion in the capital markets is well-documented,[3] but whether word-of-mouth information can lead to large-scale spreading of behaviors that can be characterized through epidemic models is an open question in the literature. In this paper I observe an epidemic solely generated by word-of-mouth communication.

[3] Shiller and Pound (1989) find that interpersonal communication is very important for individual investors' decision making. Shiller (2000, 2014) proposes that socially spreading investment information plays a crucial role in the formation of asset pricing bubbles. Empirical evidence indicates that social interaction affects bank run participation (Kelly and Ó Gráda 2000; Iyer and Puri 2012), professional money managers' portfolios (Hong, Kubik, and Stein 2005; Pool, Stoffman, and Yonker 2015), retirement plan decisions (Duflo and Saez 2002, 2003), stock market participation (Hong, Kubik, and Stein 2004; Brown, Ivković, Smith, and Weisbenner 2008) and trading behavior (Hong, Kubik, and Stein 2005; Ivković and Weisbenner 2007; Hvide and Östberg 2015). Personal information networks can affect investors' trading returns (Ozsoylev, Walden, Yavuz, and Bildik 2014) and they play an important role in illegal insider trading (Ahern 2015). See Hirshleifer and Teoh (2009) for a general review of behavior contagion in capital markets.

A major implication of the network structure I observe is that the progress of a word-of-mouth epidemic depends significantly on the distribution and structure of connections in the underlying social network, and not just on the average infection rate. The connectivity structure of scale-free networks is dominated by few highly connected hubs and models in network epidemiology show that epidemics arise and spread in scale-free networks at a much faster rate than in random spreading where each infective individual is equally likely to spread the epidemic (Pastor-Satorras and Vespignani 2001; Barthélemy, Barrat, Pastor-Satorras, and Vespignani 2004). The scale-free connectivity structure in Wincapita shows that an investment idea can spread rapidly and extensively through social interaction even if most people are just passive receivers of information, or spread the idea to only one or two others. It can also contribute to the success and survival of socially spreading Ponzi schemes, as demonstrated by the simulation findings.

More generally, my findings provide support for the use of a scale-free topology in modeling information networks among investors. Real-world social networks may not often be completely pure scale-free networks, but a power-law model clearly characterizes the Wincapita network better than a random graph model. Ozsoylev and Walden (2011) derive asset pricing implications of a scale-free information network between agents and motivate their choice of network topology by the prevalence of scale-free structures in empirical networks.

The observed S-shaped curve in information diffusion is consistent with previous hypotheses on social diffusion of interest among investors. Shiller and Pound (1989), Shiller (2000), and Shive (2010) propose that susceptible-infective-removed type epidemic models which are commonly used to model the spreading of diseases can also be used to characterize the social diffusion of interest in individual stocks. Such models imply that the cumulative number of people affected by an epidemic follows an S-shaped logistic curve over time.

The results of this paper also link to the literature on outcome-based learning in investment decisions.[4] The track record of high returns earned by the Wincapita members was undoubtedly very helpful in recruiting new investors. The relevance of peers' personally generated financial gains, together with the observed S-curve, can explain why social epidemics among investors can even take years to develop. Several generations of new investors with personal profits may be required before a socially spreading investment idea reaches a tipping point by infecting sufficiently many investors, and the growth is rapid thereafter.

The rest of the paper is organized as follows: Section II describes the details of the scheme, the data, and the characteristics of the investors, Section III studies the spreading of the scheme through network characteristics, Section IV examines the role of personal characteristics in the spreading of investment information, and Section V concludes.

## II. Background, data, and characteristics of the investors

This section provides details about the background of Wincapita and describes the data and the characteristics of Wincapita investors. The source for all information in this section (unless stated otherwise) are the police investigation documents of Wincapita.[5]

### *II.A. Wincapita as an investment scheme*

Wincapita[6] was described to its investors as an investment club that could generate significant profits on the members' investments. All operations of Wincapita took place in the internet, and

[4] Shiller (2000) proposes that extrapolation from observed high returns can generate naturally occurring Ponzi processes in the market. Kaustia and Knüpfer (2012) show that recent stock returns of peers affect stock market entry decisions. They also propose that propose that extrapolation from others' outcomes can play a part in the success of Ponzi-type securities scams. Han and Hirshleifer (2013) propose that investors are particularly likely to discuss their positive returns with others, which can results in a self-enhancing transmission bias in word-of-mouth communication.

[5] Early media reports that followed the collapse of the scheme contained many inaccuracies related to the details and rules of the scheme, because little information was publicly available. Inaccurate information can still exist in many online sources on Wincapita that use them as a reference. The police investigation, which documented the details of the scheme, was concluded in 2010 and it is the sole source of information used in this study.

the investors used the Wincapita website to manage their investment. The website could only be accessed with a personal user name and password, and it provided general information and news about the club and showed the investors how the value of their investment had developed over time. Money transfers into and out of the club were handled through a British internet payment service called Moneybookers. Wincapita had an account in Moneybookers and the investors could transfer their money by using another account in the same service. Most investors set up their own Moneybookers account, but alternatively they could transfer their money through an account belonging to a friend.

The website and the account in Moneybookers were also the entire scope of Wincapita's operations. When investors withdrew money from the scheme, it was paid out of the same account where they paid their invested funds, and no new funds were generated by actual trading or operations at any point. The virtual profits shown on Wincapita's website were completely artificial and had no link to any real-world investment assets.

The scheme was run by a single man named Hannu Kailajärvi. He had experience as a computer programmer, but no background in finance. Although some individuals helped with website updates and practical issues in different stages of the club, the police investigation indicates that Kailajärvi was the only person who knew the complete nature of the operation. Kailajärvi's identity was generally known among the investors, but he was rarely in direct personal contact with them and managed the club through the website and e-mail. The investors were lead to believe that the club was a much larger international operation. Kailajärvi e.g. used fictious names in e-mail answers to give the impression of a large number of employees and set

[6] The name of the club changed twice during its existence. Initially the club was named Giiclub, and between 2004 and 2007 it was named Winclub. For clarity, I will refer to it as Wincapita throughout the text. Wincapita is also the general name the police investigation documents use when referring to the scheme. In the scheme's internal communication, the name was often spelled with a capital C (WinCapita).

up shell companies abroad, first in Wyoming and later in Panama. In reality, the club did not even have any bookkeeping or actual legal form. The shell companies provided Wincapita with documentation that could be shown to investors, but they had no operations.

### *II.B. Timeline of the events and the end of Wincapita*

Wincapita started its operation in fall 2003 and originally the club was supposed to generate profits to its investors with a betting system for international horse racing. In early 2005 the club announced that it will shift its focus into currency trading. From that time until the end of the scheme in 2008, the source of income was claimed to be a trading system that could create large profits by day trading the EUR/USD exchange rate. In addition to the profits from currency trading, Wincapita also planned to generate additional earnings in the long run by licensing and selling its trading system to international customers. A feature that may have strengthened the investors' faith in Wincapita during the last years of the club was that they could follow trading signals allegedly generated by the club's computer system in real time through a web application. The application showed actual real time EUR/USD currency data together with different "signal markers" whose detailed functioning was not interpreted to the investors. Details about the club's operations were often described as business secrets, and the members were not provided with much information about its actual activities. They could follow the profits their investments had earned, but were not informed about the specific transactions or trades that generated them.

Wincapita first came to public attention in September 2007 when an investigative TV-journalist made a news story of the club and raised doubts about its profit mechanism speculating that it may function like a Ponzi scheme.[7] Paradoxically, many investors mention in their police

[7] The police documents identify this news story, which aired on September 23, as the first public news coverage of the scheme.

interview that the media coverage strengthened their faith in the club. The club continued its operation without any visible reaction from the authorities, and they concluded that the authorities must have investigated the club and found its operations legitimate.

The operation of the club ended in the beginning of March 2008 when Hannu Kailajärvi fled from Finland and shut down the website of the club. He hid from the police for nine months and was finally arrested in Northern Sweden in December 2008 after an international manhunt. Kailajärvi destroyed the club's records and website during the escape, and investors could not withdraw any money from the club after the website was taken down. Wincapita's bank account had a balance of 4.8 million euros at the time of Kailajärvi's escape, and the account was frozen by the authorities soon after his disappearance.

Police investigation of Wincapita's bank statements shows that the total amount of funds the investors transferred into Wincapita during its existence exceeds 100 million euros. Individual investors' police questioning transcripts indicate that in addition to the financial losses, the collapse of the club also caused significant damage through destroyed personal relationships between sponsors and their sponsored investors. The documents contain several mentions of suicides, divorces, and mental problems that resulted from the ending of the club, and the social invitation structure undoubtedly contributed to these adverse effects. The events also sparked conspiracy theories among some Wincapita members, and the interview transcripts of several investors indicate that they refused to accept that the club had been a fraud regardless of the police evidence.

The main reason why Wincapita could operate so long without interference was that it was very difficult for the Finnish authorities to receive any information on its operations. Because of the sponsoring system, there was no publicly available information about the club, and as long as no-one had suffered any losses, there was no imminent cause for a police investigation. Currency trading is less regulated than most other areas of the financial markets and the Finnish Financial Supervision Authority deemed in fall 2007 that based on available information Wincapita's operations do not fall under its supervision. Many potential sources of information were located abroad and could not be accessed by the authorities unless there was clear evidence of a crime: When Wincapita ended, its shell company was in Panama, its bank account was in the U.K. and the website was on a server in Luxemburg.

*II.C. Wincapita's rules and the incentives of the investors*

According to the police documents, the realized returns of the investors were in the range of several hundred percent over a period of six months (there is some variation depending on the time and source). Kailajärvi destroyed all Wincapita's records and web content when he fled, so the police estimates are based on witness testimonies and copies of material and printouts that were collected from the investors. Records of the virtual funds the investors had in their account were also destroyed.

When a Wincapita member made an investment, the invested funds had to stay in the club for a period of six months and could not be withdrawn before that. After six months, the funds including the generated makeshift profits could be reinvested or some or all of the funds could be withdrawn from the club. No money was paid out without a specific request. The minimum investment required for joining the club increased gradually throughout its existence, and at the end of the club it was 3,000 euros. Part of the initial investment was supposed to cover the fees of the club. After an investor had joined the club, he could invest additional funds at any time.

The members could gain additional profits by recruiting new investors into the club, but they were not required or expected to do so. The investors received 200 euros of virtual funds on their

Wincapita account for each person they sponsored, and they also received 20% of the profits generated by the funds that were invested by their sponsored investors.[8] Sponsor's sponsor would not receive any part of the profits. 10% of the generated profits were supposed to go to the club to cover its costs, so an individual investor would in total receive 70% of the virtual profits "earned" by his investment and the remaining 30% would go to the club and the sponsor.

Wincapita required that all investors who are invited to join the club must be known personally by their sponsor. The club did not want public attention, and a copy of its rules obtained by the police shows that they explicitly forbid members from distributing information about the club through websites, news groups, web forums, mass e-mails, or other forms of public media.

Although public distribution of information about the club was forbidden, active Wincapita investors occasionally organized Wincapita-related meetings that were attended by other club members and their friends who were potentially interested in investing in the club. Some of the meetings were purely social events, such as Christmas parties and boat cruises, while others were information sessions about the club and its developments. Wincapita's rules required that the meetings cannot be open to the general public and members must seek permission from the club before they organize any Wincapita-related event. The investors were asked whether they had attended any Wincapita-related meetings as part of the police interview, and 11% of the questioned investors had participated in one.[9]

[8] According to police documents, these rules were in place at least since January 1st 2007. There may have been some variation in the amounts before that. The general sponsoring system remained similar throughout the existence of Wincapita.

[9] When calculating the percentage, I count any event outside a club member's home or workplace that was attended by several club members as a meeting. Early media reports that followed the collapse of the scheme falsely claimed that most investors had joined Wincapita after attending the club's meetings or social gatherings. This percentage shows that the majority of the interviewed investors never attended any Wincapita-related event.

As Ponzi schemes usually offer significant profits to investors who join early, a natural question to ask in a Ponzi study is whether certain investors could have recognized that they are probably investing in a Ponzi scheme and tried to strategically benefit from it with the intention of reaping the early profits. In the case of Wincapita, carrying out such a strategy would have been very difficult because of Finnish legislation: Chapter 10 of Finnish Criminal Code mandates that any financial gain made as a result of criminal activity has to be paid to the state even if the person receiving the gain has not committed a crime and has acted in good faith. After the police investigation began, this law was applied to investors who had made large gains from Wincapita. An amount of their assets equal to their net gain from Wincapita was frozen by the authorities.[10] Another factor that would have made such strategic behavior very risky is that the investors were required to commit their funds for a period of six months, which is a long time to wait if the scheme can collapse at any moment. The social interaction analyses of this paper would not be affected even if some individual investors tried to behave strategically despite these considerations. The individuals would still have learned about the scheme from a familiar person and the investment decision would be based on socially transmitted information.

### *II.D. Description of data*

The dataset of this study has been hand-collected from the police investigation documents of the main criminal case on Wincapita. The police investigation material has been combined into a single formal document known as pre-trial protocol, and its official document number is 2400/R/81/10. The pre-trial protocol contains a summary of the police investigation, transcripts of questionings and interrogations carried out during the investigation, and copies of relevant

[10] The final decision on the loss of assets is made in a separate trial for each individual. These separate trials could not begin before there was a final court decision in the main case against Hannu Kailajärvi, and they are still on-going at the time of writing this paper. In a similar vein, bankruptcy clawback litigation has been used to recover false profits from Ponzi scheme investors in many cases in the U.S.

evidence material, such as bank statements, investigation reports, and e-mails. There are altogether more than 53,000 pages of material in the documents, and the majority of the pages are related to individual investors' questionings. The content of the documents is public information by court decision, but there are legal restrictions regarding the collection and use of personal information.[11]

The data on individual investors comes from the police questionings of the people who had invested money in Wincapita. The questionings took place between 2008 and 2010. Details about the data collection and specific definitions for different data items are reported in Appendix I. Data items collected from the questionings include total invested amount, first invested amount, total withdrawn amount, gender, age (at the end of 2007), coordinates of home (based on home address), education level, and a binary variable indicating whether the person is an entrepreneur. Additionally I have matched annual taxable income from the year 2007 to each investor based on a publicly available income listing.[12] The income is calculated as the sum of the person's earned and capital income. In Finnish taxation a person's taxable income is divided into these two categories, which are taxed at different rates.[13] The listing contains the earned and capital income for all tax subjects for whom either of the two exceeds 12,000 euros per year.

Sponsoring relationships are identified based on the individuals' answers to police questions. Three standard questions in the interviews were "Who was your sponsor?", "Did you sponsor anyone?" and "Do you know people who were above your sponsor in the Wincapita structure?".

[11] Finnish Personal Data Act allows the collection of personal data for scientific research purposes. Registry documentation required by the Act has been maintained throughout the data collection. The sensitive nature of the data also limits the possibilities for matching it with other data sources that contain personal information.

[12] The listing is "Veropörssi Magazine", which uses the data of the Finnish tax authority as its source. Published listings of similar scope are not available for other years.

[13] I do not separate these two sources of income in the analyses, because capital income can in many cases effectively include income from work. Many sample investors are entrepreneurs, and based on Finnish tax rules, entrepreneurs can often receive part of their work income as capital income. In such cases, the division between earned and capital income reflects tax planning and not the true nature of the income.

The relationships are identified based on these answers, and possible other information about sponsoring relationships that is mentioned elsewhere in the questioning documents. Based on the answers, I can also identify many investors who were not personally questioned by the police, and place them in the network.[14]

### *II.E. Characteristics of the investors*

Table I reports statistics of the investors. Compared to the general population, the sample investors are more educated and have higher taxable income. The percentage of investors with higher education (at least a bachelor's degree) is 37.9 and the percentage with higher than mandatory basic education 87.2. According to Statistics Finland StatFin database, the corresponding percentages among over 20-year-old Finns in 2007 were 28.3 and 69.0, so the Wincapita investors are more educated than typical Finnish adults.

Average taxable income among the investors is 47.65 thousand euros per year and the median is 33.65 thousand. The income distribution has a heavy right tail: the highest 10 percent earn over 74.4 thousand euros and there are three individuals earning more than a million euros per year. When interpreting the income figures, one should note that the observations do not include people whose income is below 12 thousand euros. The professions mentioned in the police documents indicate that many individuals who do not have income data are not in working life. People in this category include e.g. students, pensioners, and housewives. The group also includes individuals who live or work abroad.

A comparison to the general Finnish population indicates that the Wincapita investors have higher taxable income than typical Finns when accounting for the 12-thousand-euro lower limit

[14] The documents do not provide information about people who were invited to join Wincapita, but declined. The police were likely only interested in actual crime victims and Wincapita's database did not record any information about potential sponsees. As an analogy, the data I have resemble a typical dataset on disease contagion in the sense that I can observe the spreading of an epidemic, but I cannot observe people who were exposed to the epidemic and remained unaffected.

in the source data. Statistics Finland StatFin database reports the number of Finns with taxable income within different fixed intervals, and I compare Wincapita investors to the general population of Finns with taxable income over 12.5 thousand euros per year in 2007. The median category in the Finnish population is 25 to 30 thousand euros, which is below the Wincapita median of 33.65 thousand euros. The highest income category reported in the StatFin statistics is investors who earn more than 80 thousand euros per year. In the general population 2.9% of the people belong to this category, whereas among Wincapita investors the percentage is 8.8.

The taxable income figures can also include profits withdrawn from Wincapita during 2007. Investors typically reported their withdrawn funds as capital income and the tax authority taxed the income accordingly.[15] Whether Wincapita income should be considered different from other sources of income in the analyses of this paper can be debatable. It can raise the observed income above its "true" earnings-based level, but the withdrawn funds can also affect the behavior of individuals similarly as other income sources do, especially if the investors believe that Wincapita allows them to sustain a higher income level in the future. I cannot observe whether the income figures include income from Wincapita, but I run a robustness check where I exclude individuals who made withdrawals when analyzing income differences within the sponsor-sponsored investor pairs. If withdrawers are excluded, the average and median income in the data are 46.2 and 33.3 thousand euros, and the median is still higher than in the general population.

Statistics on individual investors also show that two groups of people that have been associated with overconfidence in the previous literature are well-represented in the sample: The majority of the investors are male (80%), and 26% of the investors are entrepreneurs. For comparison, the percentage of males among Finnish stock market participants is 58 (Keloharju

[15] The full tax implications of Wincapita income (had it been a real investment scheme) may have been unclear to many investors because of its international nature and opaque business model. Investors who did not report all their withdrawals as taxable income were not necessarily engaging in deliberate tax evasion.

and Lehtinen 2015) and according to OECD statistics, 6% of Finnish employed males and 2% of employed females were entrepreneurs in 2008. Males have been linked to overconfident behavior and risk-taking in investment decisions (Barber and Odean, 2001; Grinblatt and Keloharju, 2009) and entrepreneurial overconfidence has been documented e.g. by Cooper, Woo, and Dunkelberg (1988). Hvide and Panos (2014) show that stock market participation and entrepreneurship are correlated, and argue that higher risk tolerance is a possible explanation for the relationship. An interesting observation is that, at least in this particular scheme, there is no statistical evidence that a higher income or education level makes people more immune to investment scams.

Statistics on the invested amounts in Panel A of Table II show that the average amount invested in Wincapita is 15.1 thousand euros and the median is eight thousand. These invested amounts are calculated as the total amount of funds the investor transferred into Wincapita during its existence. The invested amounts are economically significant when compared to the annual income of typical investors. The median income is 34 thousand euros per year and the median investor invested an amount that corresponds to 23% of his annual income gross of taxes. Keloharju and Lehtinen (2015) report that the median Finnish stock portfolio in 2015 is worth 4,200 euros, which indicates that the median Wincapita investment is almost twice as high as the combined value of all stock holdings of a typical Finnish stock market participant. The smallest Wincapita investment in the data is 20 euros and the largest is 1.6 million euros.

The percentage of investors who withdrew money from the scheme is 26.5 and 10.6% withdrew more than 10,000 euros. The distribution of withdrawn amounts is skewed and the largest withdrawn amount is 1.7 million euros. Cases where an investor claims to have exited the scheme before its collapse are rare and account to 1.2% of the observations.

Panel B of Table II reports the distribution of investors' year of joining. Most investors joined during the last years of the scheme, 43% in 2007 and 26% in 2008 before the end of the club. Based on the annual values, the growth of Wincapita can be described as exponential. An exponential curve ($y_t = ae^{bt}$) fitted to the annual observations of the number of new investors has an $R^2$ of 90.1% whereas a linear model has an $R^2$ of 61.5%.[16]

An alternative way to measure the growth of the scheme is to study the amount of funds invested in Wincapita. Figure I shows the cumulative amount of new funds invested in the scheme between January 2005 and March 2008 based on the weekly amount of funds transferred into Wincapita's Moneybookers account. The figure starts in 2005, because it is the first year for which the invested amounts are available with weekly accuracy in the police documents. The cumulative invested amount shown in the graph grows almost linearly in logarithmic scale after the first months, which also implies that the growth of the scheme was exponential.

## III. Wincapita as a social network

In this section, I study Wincapita's sponsoring relationships as a social network. I analyze the connectivity structure of the sponsoring network, the geographic spreading of the scheme, and the cumulative number of affected investors as a function of network distance.

### *III.A. The structure of the sponsoring network and geographic spreading of the scheme*

Figure II draws a network graph of the sponsoring relationships based on all investors who can be connected to Hannu Kailajärvi through the chain of sponsors. The large plot in the middle is Hannu Kailajärvi, and the figure can be read be read as a tree diagram showing how the investment idea spread from one person to the next. In the popular press, Ponzi schemes are often depicted as pyramids, where each old investor appoints a fixed number of new members,

[16] When fitting the curve, I multiply the investors who joined in 2003 with three and investors who joined in 2008 with six to account for the fact that first investors joined in September 2003 and the scheme ended in the beginning of March 2008.

but this characterization does not fit the network structure observed in Figure II. There are large clusters of investors connected to certain sponsors, and it appears that few individuals with many connections generate the majority of the social effect.

Wincapita’s network structure also illustrates that an investor’s social influence often extends beyond the people to whom he is directly connected. If an investor in Figure II sponsors someone, in 39% of the cases at least one of the sponsored investors is also a sponsor and spreads the idea further. The household finance literature has little previous evidence on how commonly peer effects spread beyond the first step in social networks.

Figure III shows two maps that depict the geographic spreading and coverage of Wincapita. The first map illustrates the geographic network of sponsoring relationships. The lines on the map connect the locations of sponsors to the locations of their sponsored investors based on all investor pairs where both locations are known. The areas that have a large number of connections roughly correspond to the largest cities in Finland, and most geographic hubs are also connected to each other. The connections show that investment ideas travel with people, and even distant cities can have a large number of connections between them. This implies that the routes that are typically travelled by people may be a better predictor for the geographic spreading pattern of word-of-mouth information than the distance between areas.

The second map shows the geographic coverage of the scheme based on the number of sample investors as percentage of the population in Finnish municipalities. The coverage is very comprehensive and most municipalities have Wincapita investors. 18 out of the 19 Finnish administrative regions (maakunta) have observations in the data, and the only exception is the Swedish-speaking autonomous region of Åland islands.

### *III.B. Connectivity structure of the sponsoring network*

Next, I analyze the connectivity structure of the sponsoring network based on the distribution of the number of sponsored investors among sponsors (investors who invited others into Wincapita). The degree distribution, measuring the distribution of the number of connections per node is of significant interest in network analysis, because it determines the connectivity structure of a network. Among other things, it affects the flow of information in social networks (Barabási 2002).

A common observation in empirical studies on communication within social groups is that the large majority of individuals in a group obtain most of their socially transmitted information from a very small subset of the people.[17] Many empirical social networks also have a scale-free connectivity structure where the heavy right tail of the degree distribution approximately follows a power law (Barabási 2002, 2009)[18]. In contrast, if the connections between network nodes are formed randomly so that all node pairs have an equal probability being connected, the connectivity follows a Poisson distribution (Erdős and Rényi 1959).

I find that Wincapita's degree distribution is highly skewed. Panel C of Table II shows that 25% of the investors are sponsors. The average number of sponsored people among sponsors is 3.72 and the median is two. The highest number of sponsored people in the data is 120 and 2%

---

[17] Lazarsfeld, Berelson, and Gaudet (1948) and Katz and Lazarsfeld (1955) provide seminal results on this phenomenon. The fact that few socially powerful individuals, "market mavens", are very important for the diffusion of word-of-mouth information on consumer products is widely recognized in marketing (see e.g. Feick and Price, 1987). In the field of financial economics, Benartzi and Thaler (2007) provide anecdotal evidence suggesting that powerful social influencers have a strong impact on retirement plan decisions within a supermarket chain. Banerjee, Chandrasekhar, Duflo, and Jackson (2014) show that individuals can identify the influential diffusers of information in their community even if they are not aware of its social network structure. Galeotti and Goyal (2010) provide an economic model that explains the existence of few powerful influencers through the costs of acquiring personal information.

[18] The most commonly offered explanation for the power laws observed in empirical social networks is preferential attachment, where people with many social connections are more likely to generate new connections in the future. This snowball effect in connectivity over time can result in a power law distribution in the number of connections per node (Barabási and Albert 1999). A large number of other empirical regularities in economics and finance also have a power-law form (Gabaix 2009).

of the sponsors sponsored more than 20 investors. Next, I study whether the degree distribution exhibits power-law characteristics using two samples. The main sample is based on all sponsoring relationships in the data. As a robustness check, I also use another sample, which is based on the investors who contacted the police by their own initiative. I will refer to the second sample as restricted sample. (Appendix I.A. provides information on why certain people may have been contacted by the police).

The restricted sample functions as a robustness check for possible missing observations of sponsoring relationships and the possibility that there is some systematic selection bias among the investors who were approached by the police. In the restricted sample, the number of people sponsored by sponsor *i* is defined as the number of investors who approached the police themselves and identified sponsor *i* as their sponsor.

The power-law relationship $P(k) \sim k^{-\gamma}$ implies that logarithms of $k$ and logarithms of the corresponding empirical frequencies are linearly related. Figure IV draws histograms of the number of sponsored people for both samples and also draws log-log graphs of empirical frequencies and the cumulative probability distribution *P(k).* The plots in the log-log graphs for both samples approximately follow a straight line after the first observations indicating that there is a negative linear dependency between the two log-variables. The linear shapes in the graphs are consistent with a power law and the slope is somewhat steeper in the restricted sample.

If the connections between nodes in a network are formed randomly as in the Erdős-Rényi (1959) model, the connectivity always follows a Poisson distribution with a peak at some $P(\langle k \rangle)$ and a variance of $k$ equal to the mean $\langle k \rangle$. The histograms in Figure IV do not indicate such a relationship. In the main sample, mean is 3.72 and variance is 46.7 and in the second, more restricted sample mean is 1.94 and variance is 3.96. Based on these figures, the distributions are too overdispersed to fit a Poisson model, and the structure of the network is clearly different from a uniformly random model.

As a more formal test, I fit a power-law model to the data using maximum likelihood optimization. I follow the standard procedure for power-law tail analysis and fit a power-law distribution where $P(X = k) = k^{-\gamma}$ for distribution values exceeding some threshold value $k_{min}$.[19] The fitted parameters are $\gamma = 2.38$ and $k_{min} = 5$ in the main sample and $\gamma = 3.34$ and $k_{min} = 4$ in the restricted sample. The corresponding Kolmogorov-Smirnov $p$-values are 0.7139 and 0.9999. They strongly support the hypothesis that the data has a power-law tail.[20]

As a robustness check for the possibility that the observed skewness in the degree distribution can be explained by differences in investors' time of joining, I have also fit similar power-law models to subsamples consisting of investors who joined during the same year. The number of people sponsored by sponsor *i* in the subsample of year *t* is defined as the number of investors who joined Wincapita in year *t* and identified sponsor *i* as their sponsor. The power-law relationship exists also in these subsamples from six years between 2003 and 2008. The lowest $p$-value is 0.45 from year 2004 and the average is 0.86. The average $\gamma$ from the models is 2.59.

The findings demonstrate that a power-law model clearly provides a better characterization of the spreading process compared to a random graph model. Although I cannot observe the underlying network of all different types of social connections between the sample investors, the finding that the spreading process of the scheme has scale-free connectivity suggests that the relevant underlying social connections also have the same structure.

---

[19] In specific, I use power.law.fit function of package igraph in R programming language. The function is run with the default plfit implementation that finds the optimal values for $k_{min}$ and $\gamma$ using the method of Clauset, Shalizi and Newman (2009). The reason why parameter $k_{min}$ is used in empirical power-law distribution fitting is that most observed real-world power-law distributions follow the power law closely only after some threshold level (Newman 2005).

[20] In interpretation, small $p$-values (e.g. less than 0.05) indicate that the test rejects the hypothesis that the original data could have been drawn from the fitted power-law distribution.

### III.C. The actual Wincapita network compared to a simulated network

How different would the spreading process look like if the sponsoring distribution followed a Poisson model instead of the power law? I compare the actual Wincapita network shown in Figure II to a simulated random network with a Poisson distribution for the number of sponsored people. I form the simulated network so that the first sponsor is the originator of the scheme and each newly sponsored investor is a sponsor himself with probability $p$. The number of new investors sponsored by each sponsor is drawn from a Poisson distribution with mean $\lambda$. The probability $p$ is equal to the percentage of sponsors in the actual Wincapita network and $\lambda$ is the average number of sponsored people among sponsors. I continue adding nodes until the simulated network has the same number of nodes as the actual network.[21] The only difference between the actual and the simulated networks is the distribution of the number of sponsored people. Both networks have the same probability of being a sponsor ($p$) and the same average number of people sponsored by a sponsor ($\lambda$).

The shift from a power law distribution to a Poisson distribution has a dramatic effect on how the number of investors grows as a function of network distance to the originator of the scheme. Table III provides statistics comparing the actual networks to the simulated random network based on 10,000 simulation rounds. In the actual Wincapita network, all investors are within 15 steps of the starter of the scheme, whereas in the simulated network it takes on average 161 steps to reach the same number of investors. In the Wincapita network, 4,177 investors have a network distance of ten or smaller to Hannu Kailajärvi. In the simulated network, on average only 136 investors are within the same distance.

[21] If the simulation ends up in a stage where the network ceases to grow, because the most recently joined investors have not sponsored anyone, I redraw $p$ for the investors who joined in the latest period (i.e. the investors who have the longest network distance to the originator).

I also analyze how the observed network structure can contribute to the survival of Ponzi schemes. Ponzi schemes need a constant flow of money from new investors in order to finance the payouts to the existing investors. Because investors can be reached through fewer social steps in a scale-free network, the network topology can allow socially spreading Ponzi schemes to maintain a higher payout ratio without running out of money. I calibrate a simple Ponzi scheme model to the actual Wincapita network and the simulated random networks to analyze this phenomenon. In the model, the Ponzi scheme grows step by step through the network, starting from the originator of the scheme. At $t_1$ the people who are directly connected to the originator join the scheme, at $t_2$ the people who are connected to them join, and the scheme continues to grow like this until all the investors in the network have joined. All investors remain in the scheme once they have joined, so there are no exits in the model.

I assume that each investors pays the same invested amount $I$ at the time of joining and makes no further investments. Each period, the scheme pays the same fixed payment $P$ calculated as a percentage of $I$ to the investors who were members of the scheme at the end of the previous period. I calculate what is the highest value $P$ the scheme can pay the investors so that the scheme does not have a negative cash balance once it has grown through the whole network. In other words, I calculate what is the highest payout the scheme can afford to pay while remaining operational. The investors who joined in the last period will receive one payment of $P$.

I find that the actual network can sustain a payout ratio that is more than five times as high as the highest possible payout ratio in the average simulated network. The highest possible $P$ for the actual Wincapita network is 0.125, while the average value for the simulated networks is 0.024.

Altogether, the simulation results demonstrate that the observed network topology can contribute to the growth and survival of Ponzi schemes.

*III.D. Implications of the scale-free connectivity structure*

The scale-free network literature offers a formal explanation for why information spreads significantly faster in networks that have a power-law degree distribution. Because of the power law, the structure of all scale-free networks is dominated by few highly connected hubs, and by construction scale-free networks have very short average node to node distances (Cohen and Havlin 2003). Models in network epidemiology show that epidemics spread through scale-free networks at a much faster rate than in models of random spreading where each infective individual is equally likely to spread the epidemic (Pastor-Satorras and Vespignani 2001; Barthélemy, Barrat, Pastor-Satorras, and Vespignani 2004). An epidemic starting at a random point in any scale-free network will quickly reach a highly connected hub, and in the next stage the hub will infect a large number of nodes. In other words, people with many social connections facilitate the spreading of an epidemic because of the joint effect of two factors: Their high number of social connections means that they are more likely to be infected early, and once they are infected they will next spread the epidemic to a large number of people who are within one social step in the network.

In line with this idea, most investors in the Wincapita network are directly connected only to their sponsor, but at the same time, they are typically only a short network distance away from a highly connected network hub, which connects them to a much larger number of people. The average investor in the sponsoring network of Figure II is connected to only 2.0 other investors, but the number of people within two steps is 20.5 and the number of people within three steps is

76.5. These figures are much higher than, for example, in a network where every person has connections to exactly two other people in addition to their sponsor.

Consistent with models of epidemic spreading in scale-free networks, I also find that highly connected nodes are reached in the early stages of the Wincapita epidemic. Table IV divides investors into groups based on fixed network distance intervals and network distance quartiles using the number of sponsors between the investor and the starter of the scheme as network distance. The average number of people sponsored by an investor in a group decreases monotonically with the network distance.

Another important implication of the scale-free network structure is that even weakly contagious word-of-mouth epidemics can spread widely in the population. Traditional susceptible-infective-removed type epidemic models, such as those referred to in Shiller's famous book *Irrational Exuberance* (2000), predict a critical threshold for the propagation of an epidemic throughout a population. If a disease is less infectious than that epidemic threshold, the epidemic will die out, whereas epidemics where the spreading rate is above the threshold will multiply exponentially, and penetrate the entire system. Pastor-Satorras and Vespignani (2001) show that in scale-free networks the epidemic threshold always converges to zero. That is, even weakly contagious epidemics will spread widely and persist in scale-free networks. This well-known property has since been used to explain the rapid spreading and persistence of sexually transmitted diseases and computer viruses (Barabási 2009).[22]

As a result of this property, the scale-free connectivity structure in Wincapita provides a network theory-based explanation for why socially transmitted investment ideas can spread very rapidly even if the average person discusses them with only few other people. The structure also

[22] This follows from the empirical observation that human sexual contacts and computer networks also have a scale-free network structure (Barabási 2009).

explains why Ponzi schemes and other asset-specific fads can survive and grow for a long time even if most people are not gullible to the investment idea.

The fast spreading rate of Wincapita is difficult to explain without accounting for the network structure. On average it takes a relatively long time for the idea to be transmitted from one person to the next, and the average number of new investors invited by an existing member combined with the hazard rate cannot alone generate the rapid spreading observed in the data. The data has 1,248 sponsor - sponsored investor pairs where I can identify the month of joining of both the investor and the sponsor. The average difference in their time of joining is 10.7 months and the median is 7. At the average rate of 3.7 investors sponsored by a sponsor, the scheme would never grow to cover thousands of investors in less than five years if the spreading process followed a Poisson topology.

Finally, one can ask to what extent are the network topology findings of this paper generalizable to other settings? Wincapita participants may differ from typical retail investors in many aspects, but a growing body of literature shows that social networks are structurally similar in many different contexts and among different cultural and sociodemographic groups around the world. At least on this basis, there is no cause to believe that the social connections of Wincapita participants are fundamentally different from the connections among other types of people.

*III.E. Network distance and the spreading dynamics of Wincapita*

I also use the Wincapita network to characterize how the aggregate number of investors affected by an investment idea grows as a function of social distance from the originator of the idea. This analysis provides information about the spreading dynamics of word-of-mouth information within social networks. Previous research on social networks suggests that even random people unknown to each other are typically connected by only few social network links.

In his classical study of the "six degrees of separation" Milgram (1967) finds that people are on average six social contacts away from any random person, and the average path lengths between random people display strong regularities. This "small world" phenomenon has since been explained with the common structures of social networks (Watts and Strogatz 1998).

Figure V shows graphs depicting the cumulative number of investors in Wincapita as a function of network distance. The network distance of investor *i* is calculated as the number of sponsors between investor *i* and Hannu Kailajärvi in the Wincapita network shown in Figure II. The curves measure how the size of the scheme grows if it starts at Hannu Kailajärvi and spreads in the network of Figure II by one step at a time. The upper graph in Figure V is drawn based on all sample investors and the lower graph draws the same curve based on two subsamples of investors who joined Wincapita between 2003-2006 and 2007-2008.

All three curves are S-shaped, indicating that the diffusion of information progresses in a nonlinear fashion, and there is strong regularity in the typical network distance. In the full sample, the median distance to Kailajärvi is eight steps and the growth rate slows down after nine steps. In the 2003-2006 subsample the median distance is seven steps (maximum is 14) and in the 2007-2008 subsample the median is nine steps (maximum is 15). The growth rate slows down after eight steps in the 2003-2006 sample and after nine steps in the 2007-2008 sample. The fact that the maximum and median network distances in both subsamples are very close to each other indicates that the typical network distance to Kailajärvi does not grow at the same rate as the number of investors, which is consistent with the "small world" phenomenon.

The S-shaped curve is interesting also, because it is in line with the predictions of epidemic models of investor behavior proposed by Shiller and Pound (1989), Shiller (2000), and Shive (2010). In these models, the number of people affected by an idea or behavior grows approximately exponentially in the early stages of the social epidemic, and the growth rate slows down over time, because there are fewer people left who are willing to adopt the idea but have not already done so. The models imply an S-shaped pattern in information diffusion. Shiller (2000) proposes that the S-curve can be used to characterize the social transmission of attitudes in speculative bubbles.

There are at least two potential explanations for why the growth rate of Wincapita slows down after a number of social steps. One possibility is that people who are late joiners find it more difficult to recruit new investors, because many of their friends who are potentially interested in the scheme have already joined. The scale-free epidemic literature offers a complementary network theory-based explanation. If social networks are scale-free, late joiners are more likely to be people with few social connections and fewer potential friends to invite.

## IV. Personal characteristics and the spreading of word-of-mouth information

In this section, I study how personal characteristics are related to the spreading of Wincapita at the level of individual investors. Despite the extending literature on peer effects in financial decision-making, there is little previous field evidence on whose investment ideas investors typically follow. Personal characteristics can be correlated with a person’s willingness to distribute information about Wincapita and with the ability to persuade others to invest in it. I cannot separate these two channels in the analyses, because I do not have data on people who were invited to join Wincapita, but declined. However, the results indicate the direction in which the information typically flows among people with different characteristics. The findings also provide statistical evidence about the characteristics associated with the influential sources of information.

The police interview transcripts show that Wincapita spread through many different kinds of social connections. Sponsors include friends, acquaintances, co-workers, business partners, relatives, neighbors, and people known through hobbies and religious communities. The police did not explicitly ask about the nature of the social relationship to the sponsor, but 15.5% of the questioned investors mention that their sponsor was a relative or family member and 8.2% mention that their sponsor was a co-worker.

### *IV.A. Sponsors compared to non-sponsors*

Table V compares the personal characteristics of sponsors and non-sponsors and also reports statistics on the characteristics of the most active sponsors in the sample. The statistics are based on the investors who were personally interviewed by the police. To measure whether the difference between sponsors and non-sponsors is statistically significant, I calculate simulation-based *p*-values by assigning the same number of sample investors randomly as sponsors. The two-sided *p*-value is the probability that the difference between randomly assigned sponsors and non-sponsors is at least as large as the actual difference based on 1,000 simulation rounds.

Table V shows that sponsors typically have comparatively higher income and are more likely to be male. The difference in average annual income is 7.4 thousand euros, and the percentage of females is nine percentage points lower among non-sponsors. Both differences are statistically significant. The difference in income is even more pronounced when comparing the most active sponsors with non-sponsors. The average income among the investors who sponsored at least ten people is more than twice as high as among non-sponsors. The average sponsor is slightly younger than the average non-sponsor, but the people who sponsored at least ten investors are on average older that non-sponsors.

For legal and ethical reasons, I cannot provide very detailed personal information about the most active sponsors, but the statistics on Table V show that they are also typically males with high income. Based on the police interviews, most of them have an occupation or position, which puts them in contact with a large number of people on a daily basis. Examples include business-to-business sales representatives, corporate managers, and people who are active within religious communities. This is consistent with the idea that people with a large pre-existing social network were the most powerful spreaders of Wincapita.

### *IV.B. Differences and similarities within the sponsor-sponsored investor pairs*

Next, I analyze differences in personal characteristics within individual sponsor-sponsored investor pairs. A very common phenomenon in social networks is homophily, the tendency of individuals to form ties with others who share similar sociodemographic or personal characteristics (McPherson, Smith-Lovin, and Cook, 2001). It is therefore natural to expect that the sponsors and their sponsored investors are on average very similar in terms of different personal characteristics. Systematic differences within the pairs are interesting, because they can reveal which characteristics are relevant to the transmission of word-of-mouth information.

Table VI reports differences in characteristics based on age, education, gender, geographic distance, and income. For each statistic, I also report the corresponding value from a simulation where each investor is given a random place in the network without replacement. I run the simulation separately for each data item, so that only the investors with a non-missing value for the item change places. The reported simulated statistics are average values based on 1,000 repetitions, and the table also reports the probability of getting a statistic that is comparable to the empirically observed value. I use the simulation to measure statistical significance, because it accounts for the topology of the network and the empirical distributions of the variables.

The results show that the sponsor has typically comparatively higher income and is slightly older. Males typically match with other males, whereas among females, the gender of the sponsor is not statistically different from random. The median difference in income is 9.5 thousand euros per year, and the average difference is 65 thousand. The sponsor has higher income in 61% of the pairs. There is strong skewness in the differences, and the third quartile of the income difference is 73 thousand euros. The interview transcripts indicate that there are several cases where an entrepreneur or executive sponsored many of their lower-level employees, which is one possible source for the largest differences in the data. If sponsors who withdrew money from Wincapita are excluded, the average difference is still 56.2 thousand euros and a difference of the same magnitude is not obtained in any simulation round.

The median age difference is one year and the average is 1.4 years. Values higher than the average and median are produced in 9% of the simulation rounds, so both the average and the median are statistically significant at the 10% level. Based on the year of birth, the sponsor is older in 51.2% of the observations and younger in 43.4% of the observations. Although the percentage of older sponsors is only slightly over 50%, the empirical probability of matching with an older sponsor is still almost 8 percentage points higher than the probability of matching with a younger sponsor. There is also evidence that age cohorts may matter in the diffusion of information. The probability of matching with a sponsor who is born in the same year is 5.4%, and the simulation does not produce a similar or higher percentage in any round.

There are no statistically significant differences in education, based on three education level categories identifying investors with mandatory basic education, upper secondary education, and higher education. The investor and the sponsor have the same education level in 53% of the pairs, which is higher than in any of the simulation rounds. The empirical probability of having a lower education sponsor exceeds the probability of having a higher education sponsor (25.9 vs. 21.4%), but the difference is not statistically significant. Larger differences between the two categories are obtained in 17.2% of the simulation rounds.

As can be expected, geographic distance in most sponsoring relationships is short. The median distance to sponsor is 16.1 kilometers, and the average distance is 76.4 kilometers. In 9.8% of the observations the sponsor lives within one kilometer, which suggests that many of the sponsors can be classified as neighbors. Previous studies that identify neighbors based on zip codes have found that neighbors affect stock market participation decisions (Brown, Ivković, Smith, and Weisbenner 2008; Kaustia and Knüpfer 2012).

I also find that the correlation in personal characteristics often extends beyond the first step in the social network. Table VII reports pairwise Pearson correlations in personal characteristics between an investor and a person who is connected to him through the chain of sponsors and *n* steps closer to the starter of the scheme. E.g. the person with a network distance of zero to the investor is his sponsor and the person with a network distance of one is the sponsor's sponsor. The personal characteristic variables are age, income, a binary variable for females, and a binary variable for investors with higher education. With continuous variables, I exclude the highest and lowest 1% of the variable values to limit the effect of outliers. Based on the coefficients, investors within two social steps are statistically significantly correlated with all variables except income, which is only statistically significant within the first step.

### *IV.C. Implications of the findings*

Overall, the findings in this Section show that some personal characteristics are relevant to the spreading of investment information. Age and income are potential sources of credibility, which can be one explanation for why these characteristics are associated with sources of word-of-mouth information. Evidence from a field experiment by Bursztyn et al. (2014) indicates that social learning effects in financial decision-making are stronger when investors observe the decision of a financially sophisticated peer.

Another possible explanation for why most people joined through someone with comparatively higher income can be a “keeping up with the Joneses” effect where investors invest in the same asset as their peers because of relative wealth concerns (Abel, 1990; Gali, 1994; Bakshi and Chen 1996; DeMarzo, Kaniel, and Kremer 2008). Wincapita was a get-rich-quick scheme, and investors may have been worried about losing in wealth relative to their peer if they do not invest similarly. Prospect theory suggests that if peers’ wealth is a reference point, the difference relative to a perceivably wealthier peer can be particularly relevant for investment decisions. Taxable income is not a direct measure of wealth, but it is likely correlated with the person’s perceived financial status and consumption.

On average, most investor pairs are close to each other in terms of personal characteristics, and the correlation coefficients show that the similarity extends to people who are within two steps in the social network. While this finding is not very surprising, the homophily in social interactions can nevertheless be relevant for explaining differences in investment behavior across social and demographic groups. Word-of-mouth communication that mainly takes place within homogeneous social groups, rather than across them, can strengthen homogeneity in economic behavior within the groups (Granovetter 2005) and preclude or slow down the convergence of behaviors in the population (Golub and Jackson 2012). It can also significantly facilitate the diffusion of new ideas within social groups where members share characteristics that make a particular idea more attractive to them (Jackson 2014).

## V. Conclusion

The Wincapita dataset provides information about the social process through which investment ideas spread from one person to the next. The findings cannot be characterized by models of information diffusion in which investment ideas are transmitted evenly from one investor to a fixed number of peers. Instead, a small fraction of the investors is responsible for the majority of the observed social effect, and most investors are passive receivers who adopt the idea but do not spread it further. The findings suggest that social network structures play an important role in the spreading of investment ideas and few powerful individuals with many social connections can significantly facilitate the epidemic spreading of contagious ideas.

The observations of this paper also contribute to the discussion on the benefits and harms of peer effects in financial decision-making. Although investors' social learning can produce welfare-improving outcomes in many situations e.g. through higher stock market participation and better portfolio diversification, the evidence of this paper indicates that it can also spread and exacerbate investment mistakes. One potential explanation for the social spreading of investment mistakes observed in this study is a heuristic known as question substitution. When faced with a difficult question, people often answer an easier one instead (Kahneman 2011). In the case of Wincapita some investors may have exchanged the more complex question "Do I trust this investment scheme?" to the simpler question "Do I trust the person who is telling me about this investment scheme?"

**Appendix I. Details on the police interview process and data collection**

This appendix provides details about the police interview process and explains how the data items were collected from the police investigation material.

**A. The police interviews**

The interviews were usually carried out in a police station close to the investor's home. In individual cases they were carried out over the telephone or e-mail, if it was very difficult to reach the individual. The interviewed persons were typically asked the same questions in the same order. The questioning began with general questions about the investor's actions in Wincapita, followed by detailed questions about the investment and sponsoring activities. The interview documents contain full transcripts of the individuals' answers and the interviewing police officer's questions and they were signed by the questioned person at the end of the interview.

Most interviewed investors (57%) contacted the police by their own initiative and 43% were invited to the interview by the police. The interview documents do not reveal why a particular person was approached by the police, but potential reasons include at least large money withdrawals from Wincapita and an active role in the club. Some questioned investors were asked to verify other investors' earlier statements, which may also have been one reason to contact specific individuals. Naturally, it is possible that some of the investors who were called to an interview would have approached the police by their own initiative if the police had not contacted them first.

Investors who had lost money could make claims based on their net loss (total amount invested minus total amount withdrawn) and the public prosecutor would drive these claims as part of the main Wincapita case against Kailajärvi. The police made public announcements asking investors to contact them to claim losses and to provide information about the investigated crime. All investors who contacted the police were questioned to verify their activities in Wincapita and the amounts they had invested. If an investor claimed losses, he had to provide evidence that the claimed amounts had been transferred into or out of Wincapita. Typically, such evidence consisted of bank statements or receipts. The police had access to the bank statements of Wincapita's Moneybookers account, and they were the main source of information when money transfers were verified. Even though the police did not interview all the individual investors, they were able to estimate the total number of Wincapita participants based on the bank account information and other sources, such as confiscated computer files that contained some of the scheme's records from the year 2007.

Most investors were co-operative throughout the police interview, but there are individual investors who refused to provide information or answer specific questions. In such cases, the individuals usually refused to believe that Wincapita had been an illegal investment operation and mentioned this to the interviewer. The police recorded the identity of all the persons who were interviewed as part of the police investigation.

**B. The collection of data items**

I have collected all statistical information about Wincapita investments and the demographic characteristics of the investors that is generally available in the documents. The recorded personal characteristic items are age, education, gender, home address, profession, and a binary variable indicating whether the person is an entrepreneur. The first page of the interview document is always a standard police form, which identifies the questioned person and records personal background and contact information. I use it to identify the person's age, profession, and home address. In the beginning of the interview, the interviewed person was also asked to describe his education and professional background. I use the answer to this question to define the education level and to complement the information on the self-reported profession. Information about invested and withdrawn amounts and the investors' time of joining Wincapita are based on answers to specific police investigation questions. I provide details about how individual data items were collected below.

**Age**

Age is based on the date of birth. It is calculated at the end of 2007, which is the last full year of Wincapita's operations. I obtain the date of birth from the first six digits of the Finnish social security number on the police form. The police recorded the date of birth separately in case the person did not have a social security number.

**Education**

I measure education based on the highest degree possessed by the person. I divide the degrees into five different categories, which are mandatory education, upper secondary education, higher education (bachelor's degree or similar), master's degree, and doctoral degree. In the statistical analyses of the paper, I typically combine the three highest categories into a single "higher education" category. The Statistics Finland StatFin database uses the same three-level classification. Finnish degree titles that correspond to different education levels are governed by law, so a person's level of education can be identified accurately if the degree title is mentioned in the interview. The availability of the education item is slightly lower than the availability of other data items, because the degree title is not always mentioned in the verbal answer.

**Gender**

I identify the person's gender based on the first name or the last digit of the social security number, if it is recorded. In Finland, it is typically possible to identify a person's gender based on the first name. Finnish legislation mandates that males are not allowed to have first names that are generally associated with females and vice versa. I also use the first name to record the gender of the investors who were not personally interviewed by the police, if it is possible. The last digit of the Finnish social security number provides an alternative way to identify the gender of the interviewed investors. It is an odd number for males and an even number for females.

**Home address and the coordinates of home**

The home address is the address reported in the police investigation document. It records the home location at the time of the interview. I use the coordinates of the address to calculate distances between investors and their sponsors. The distances are calculated using the haversine formula, which accounts for the curvature of the earth.

**Profession**

I collect the self-reported profession from the police information sheet. I use the self-reported profession to identify entrepreneurs.

**Entrepreneurs**

I classify an individual as an entrepreneur if he has mentioned “entrepreneur” as his profession, or otherwise mentioned that he owns a private business or company. I do not have any detailed information about the nature of the business, and investors are classified as entrepreneurs even if the private business is not their primary source of income.

**Information about Wincapita investments and withdrawals**

The invested and withdrawn amounts of money and the time of joining are recorded based on interview answers. The investors were asked separately how much money they had invested in Wincapita and how much money they had withdrawn from Wincapita. The interviewing officer compared this information to the information the police had, such as Wincapita’s Moneybookers statements and other bank information. The detailed account statements of Wincapita's Moneybookers account are not included in the pre-trial protocol, but they are frequently referred to in the interview documents.

As the investors could make Wincapita-related money transfers through other people's Moneybookers accounts, a person's money transfers into and out of Wincapita's Moneybookers account were not always related to his personal investments. One purpose of the interviews was to identify what the exact amounts personally invested and withdrawn by each investor were. The police documents do not typically contain the exact timing of each separate cash flow. The first invested amount can be identified in most cases, because each investor was asked to explain how and when they had joined Wincapita. Investors became members of Wincapita after depositing the initial investment, and the amount is usually mentioned in the interview or the accompanying documents. The year of joining and month of joining (if available) are recorded based on the same answer.

In individual cases where the investors had some uncertainty about the exact amount or could not provide bank statements to prove the transactions, the amounts are recorded based on the investor's estimate, unless the police officer had conflicting information. In the uncertain cases the individual had usually given cash to another Wincapita member who transferred the funds into the club's account. In case the interviewed person refused to provide information about the investments and withdrawals, the amounts are recorded based on the amounts the police officer had presented to the individual as his likely investments and withdrawals.

## References


Abel, Andrew. B., "Asset prices Under Habit Formation and Catching Up with the Joneses," *American Economic Review,* 80 (1990), 38–42.

Ahern, Kenneth R., "Information Networks: Evidence from Illegal Trading Tips," *Journal of Financial Economics,* Forthcoming, 2015.

Albert Réka, and Barabási, Albert-László, "Statistical Mechanics of Complex Networks," *Reviews of Modern Physics*, 74 (2002), 47-97.

Andrei, Daniel, and Cujean, Julien, "Information Percolation, Momentum, and Reversal," Working paper, 2015.

Bakshi, Gurdip S. and Chen, Zhiwu, "The Spirit of Capitalism and Stock Market Prices," *American Economic Review*, 86 (1996), 133–157.

Banerjee, Abhijit, Chandrasekhar, Arun G., Duflo, Esther, and Jackson, Matthew O., "Gossip: Identifying Central Individuals in a Social Network," Working paper, 2014.

Banerjee, Abhijit and Fudenberg, Drew, "Word-of-Mouth Learning," *Games and Economic Behavior,* 46 (2004), 1-22.

Barabási, Albert-László and Albert, Réka "Emergence of Scaling in Random Networks," *Science*, 286 (1999), 509-512.

Barabási, Albert-László. *Linked: The New Science of Networks*. (Cambridge, MA: Perseus Publishing).

Barabási, Albert- László, "Scale-Free Networks: A Decade and Beyond," *Science,* 325 (2009), 412-413.

Barber, Brad M., and Odean, Terrance, "Boys will be Boys: Gender, Overconfidence, and Common Stock Investment," *Quarterly Journal of Economics,* 116 (2001), 261-292.

Barthélemy, Marc, Barrat, Alain, Pastor-Satorras, Romulado, and Vespignani, Alessandro, "Velocity and Hierarchical Spread of Epidemic Outbreaks in Scale-Free Networks," *Physical Review Letters*, 92 (2004), 178701.

Benartzi, Shlomo and Thaler, Richard H., "Heuristics and Biases in Retirement Savings Behavior," *Journal of Economic Perspectives,* 21 (2007), 81-104.

Brown, Jeffrey R., Ivković, Zoran, Smith, PaulA., Weisbenner Scott J., "Neighbors Matter: Causal Community Effects and Stock Market Participation," *Journal of Finance,* 63 (2008), 1509-1531.

Bursztyn, Leonardo, Ederer, Florian, Ferman, Bruno, and Yuchtman, Noam, "Understanding Mechanisms Underlying Peer Effects: Evidence from a Field Experiment on Financial Decisions," *Econometrica*, 82 (2014), 1273-1301.

Cao, H. Henry, Han, Bing and Hirshleifer, David. "Taking the Road Less Traveled by: Does Conversation Eradicate Pernicious Cascades?" *Journal of Economic Theory* 146 (2011), 1418-1436.

Cohen, Reuven, and Havlin, Shlomo, "Scale-Free Networks Are Ultrasmall," *Physical Review Letters*, 90 (2003), 058701.

Clauset, Aaron, Shalizi Cosma R., and Newman Mark E. J., "Power-Law Distributions in Empirical Data," *SIAM Review,* 51 (2009), 661-703.

Cooper, Arnold C., Woo, Carolyn Y., and Dunkelberg William C., "Entrepreneurs' Perceived Chances for Success," *Journal of Business Venturing,* 3 (1988), 97-108.

DeMarzo, Peter M., Kaniel, Ron, and Kremer, Ilan, "Relative Wealth Concerns and Financial Bubbles," *Review of Financial Studies,* 21 (2008), 19-50.

Duflo, Esther, and Saez, Emmanuel, "Participation and Investment Decisions in a Retirement Plan: The Influence of Colleagues' Choices," *Journal of Public Economics,* 85 (2002), 121-148.

Duflo, Esther, and Saez, Emmanuel, "The Role of Information and Social Interactions in Retirement Plan Decisions: Evidence from a Randomized Experiment," *Quarterly Journal of Economics,* 118 (2003), 815-842.

Ellison, Glenn, and Fudenberg, Drew, "Word-of-Mouth Communication and Social Learning," *Quarterly Journal of Economics,* 110 (1995)., 93-125.

Erdős, Paul, and Rényi, Alfréd, "On Random Graphs," *Publicationes Mathematicae*, 6 (1959), 290–297.

Esitutkintapöytäkirja 2400/R/61/10 (2010). Finnish National Bureau of Investigations. (The police investigation documents of Wincapita)

Feick, Lawrence F., and Price Linda L., "The Market Maven: A Diffuser of Marketplace Information," *Journal of Marketing,* 51 (1987), 83-97.

Fruchterman, Thomas M. J., and Reingold, Edward. M. "Graph Drawing by Force-directed Placement," *Software - Practice and Experience*, 21 (1991), 1129-1164

Gabaix, Xavier, "Power Laws in Economics and Finance," *Annual Review of Economics*, 1 (2009), 255-293.

Galeotti, Andrea, and Goyal, Sanjeev, "The Law of the Few," *American Economic Review,* 100 (2010)., 1468-1492.

Golub, Benjamin, and Jackson, Matthew, O. "How Homophily Affects the Speed of Learning and Best-Response Dynamics," *Quarterly Journal of Economics,* 127 (2012), 1287-1338.

Granovetter, Mark, "The Impact of Social Structure on Economic Outcomes," *Journal of Economic Perspectives,* 19 (2005), 33-50.

Grinblatt, Mark, and Keloharju, Matti, "Sensation Seeking, Overconfidence, and Trading Activity," *Journal of Finance,* 64 (2009), 549-578.

Han, Bing, and Hirshleifer, David, "Self-Enhancing Transmission Bias and Active Investing," Working Paper, 2013.

Han, Bing, and Yang, Liyan, "Social Networks, Information Acquisition, and Asset Prices," *Management Science,* 59 (2013), 1444-1457.

Hirshleifer, David, and Teoh, Siew Hong "Thought and Behavior Contagion in Capital Markets" In *Handbook of Financial Markets: Dynamics and Evolution, Thorsten Hens and Klaus Reiner Schenk-Hoppé* Eds. (Elsevier/North-Holland, 2009)

Hong, Harrison, Kubik, Jeffrey D., and Stein, Jeremy C., "Social Interaction and Stock-Market Participation," *Journal of Finance,* 59 (2004), 137-163.

Hong, Harrison, Kubik, Jeffrey D., and Stein, Jeremy C., "Thy Neighbor's Portfolio: Word-of-Mouth Effects in the Holdings and Trades of Money Managers," *Journal of Finance,* 60 (2005), 2801-2824.

Hvide, Hans K., and Panos Georgios A., "Risk Tolerance and Entrepreneurship," *Journal of Financial Economics,* 111 (2014), 200-223.

Hvide, Hans K., and Östberg, Per, "Social Interaction at Work," *Journal of Financial Economics,* 117 (2015), 628-652.

Ivković, Zoran, and Weisbenner Scott J. "Information Diffusion Effects in Individual Investors' Common Stock Purchases: Covet Thy Neighbors' Investment choices," *Review of Financial Studies,* 20 (2007), 1327-1357.

Iyer, Rajkamal, and Puri, Manju, "Understanding Bank Runs: The Importance of Depositor-Bank Relationships and Networks," *American Economic Review,* 102 (2012), 1414-1445.

Jackson, Matthew O., "Networks in the Understanding of Economic Behaviors," *Journal of Economic Perspectives,* 28 (2014), 3-22.

Jackson, Matthew O., and Rogers, Brian W., "Meeting Strangers and Friends of Friends: How Random Are Social Networks?" *American Economic Review,* 97 (2007), 890-915.

Jackson, Matthew O., Rogers, Brian W., and Zenou Yves, "The Economic Consequences of Social Network Structure," Working paper, 2015.

Kahneman, Daniel, *Thinking, Fast and Slow,* (New York: Farrar, Straus and Giroux).

Katz, Elihu, and Lazarsfeld, Paul Felix, *Personal Influence: The Part Played by People in the Flow of Mass Communications,* (New York: The Free Press).

Kelly, Morgan, and Ó Gráda, Cormac, "Market Contagion: Evidence from the Panics of 1854 and 1857," *American Economic Review,* 90 (2000), 1110–24.

Kaustia, Markku, and Knüpfer, Samuli, "Peer Performance and Stock Market Entry," *Journal of Financial Economics,* 104 (2012), 321-338.

Keloharju, Matti, and Lehtinen, Antti, "Shareownership in Finland 2015," *Nordic Journal of Business,* 64 (2015), 182-206.

Lazarsfeld, Paul, Berelson, Bernard., and Gaudet, Hazel, *The People's Choice: How the Voter Makes Up His Mind in a Presidential Campaign*, (New York: Columbia University Press).

McPherson, Miller, Smith-Lovin, Lynn, and Cook, James. M., "Birds of a Feather: Homophily in Social Networks," *Annual Review of Sociology,* 27 (2001), 415-444.

Milgram, Stanley, "*The Small-World Problem," Psychology Today, 2 (1967), 60–67.*

Newman, Mark E. J., "Power Laws, Pareto Distributions and Zipf's Law," *Contemporary Physics,* 46 (2005), 323-351.

Ozsoylev, Han N., and Walden John, "Asset Pricing in Large Information Networks," *Journal of Economic Theory,* 146 (2011), 2252-2280.

Ozsoylev, Han. N., Walden John, Yavuz, M. Deniz, and Bildik, Recep, “Investor Networks in the Stock Market,” *Review of Financial Studies* 27 (2014), 1323-1366.

Pastor-Satorras, Romualdo, Vespignani, Alessandro, “Epidemic Spreading in Scale-Free Networks,” *Physical Review Letters*, 86 (2001), 3200-3202.

Pool, Veronika, K., Stoffman, Noah, and Yonker, Scott, E., “The People in Your Neighborhood: Social Interactions and Mutual Fund Portfolios,“ *Journal of Finance* 70 (2015) 2679–2732.

Shiller, Robert J., and Pound, John, “Survey Evidence on Diffusion of Interest and Information Among Investors,” *Journal of Economic Behavior and Organization*, 12 (1989), 47-66.

Shiller, Robert J., *Irrational Exuberance*, (New Jersey: Princeton University Press).

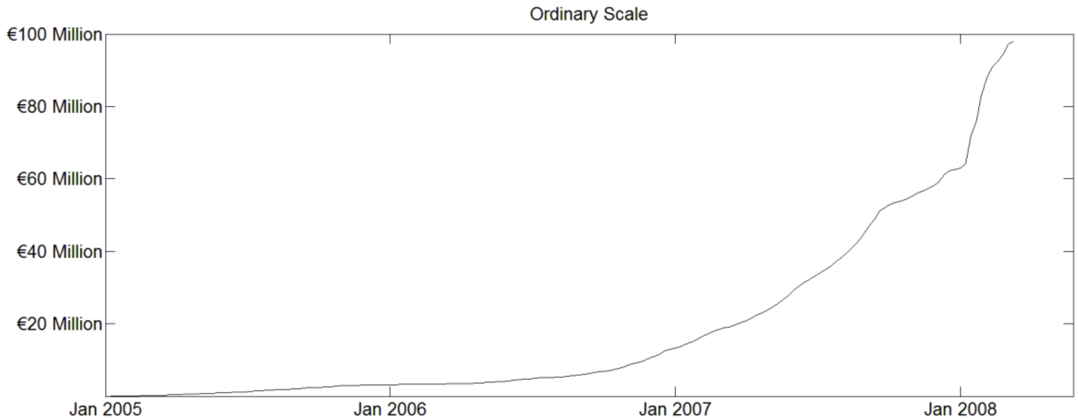
Shiller, Robert J., “Speculative Asset Prices,” *American Economic Review*, 104 (2014), 1486-1517.

Shive, Sophie, “An Epidemic Model of Investor Behavior,” *Journal of Financial and Quantitative Analysis*, 45 (2010), 169–198.

Stein, Jeremy C., “Conversations Among Competitors,” *American Economic Review*, 98 (2008), 2150-2162.

Watts, Duncan J., and Strogatz, Steven H., ”Collective Dynamics of 'Small-World' Networks,” *Nature*, 393 (1998), 440-442.

YLE News, ”WinCapitasta tulossa maan laajimpia rikostutkintoja” http://yle.fi/uutiset/wincapitasta_tulossa_maan_laajimpia_rikostutkintoja/5836796, (2008)



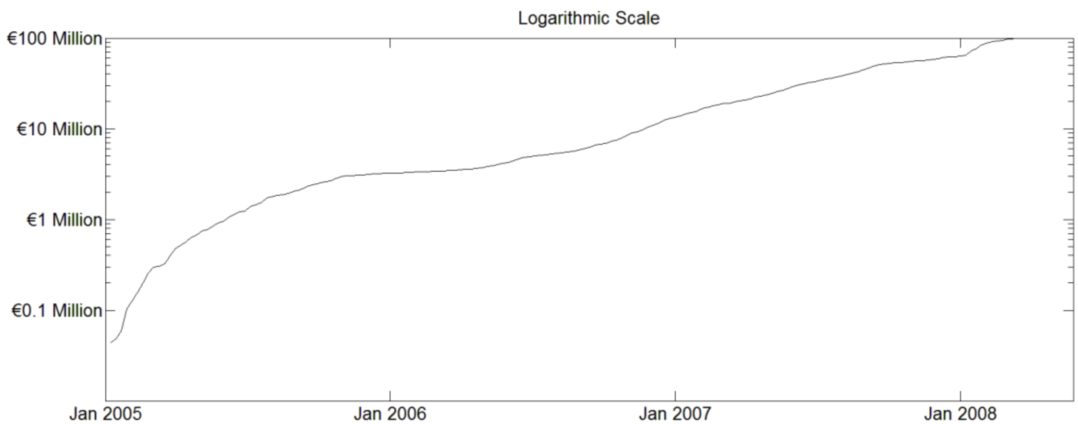




**Figure I**

**The Cumulative Amount of Funds Invested in Wincapita since January 1st 2005**

This figure shows the cumulative amount of funds invested in Wincapita between the beginning of 2005 and the end of the club in 2008. The figure is based on weekly observations of the amount of new funds transferred into Wincapita's account. The amounts invested before 2005 are not available with weekly accuracy in the policy documents and are therefore not included in the figure. The y-axis has ordinary scale in the upper graph and logarithmic scale in the lower graph.

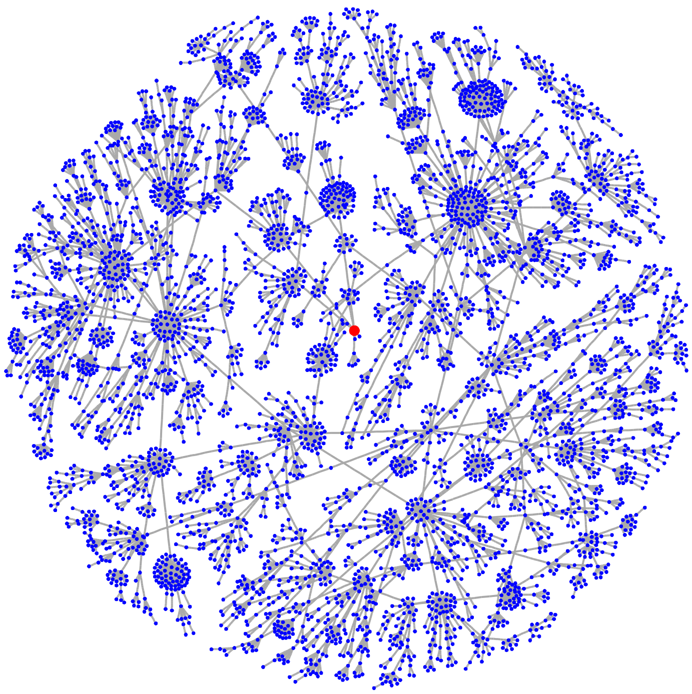


**Figure II**
**Wincapita's Sponsoring Relationships as a Social Network**

This graph illustrates Wincapita as a social network. The plots are individual investors and the lines connecting them represent the sponsoring relationships. The larger plot in the middle is the originator of the scheme, Hannu Kailajärvi. The graph includes all investors who can be linked to Hannu Kailajärvi through the chain of sponsors. It has been drawn using the Fruchterman-Reingold (1991) algorithm. The distances between points are determined by the algorithm and have no economic interpretation as such.

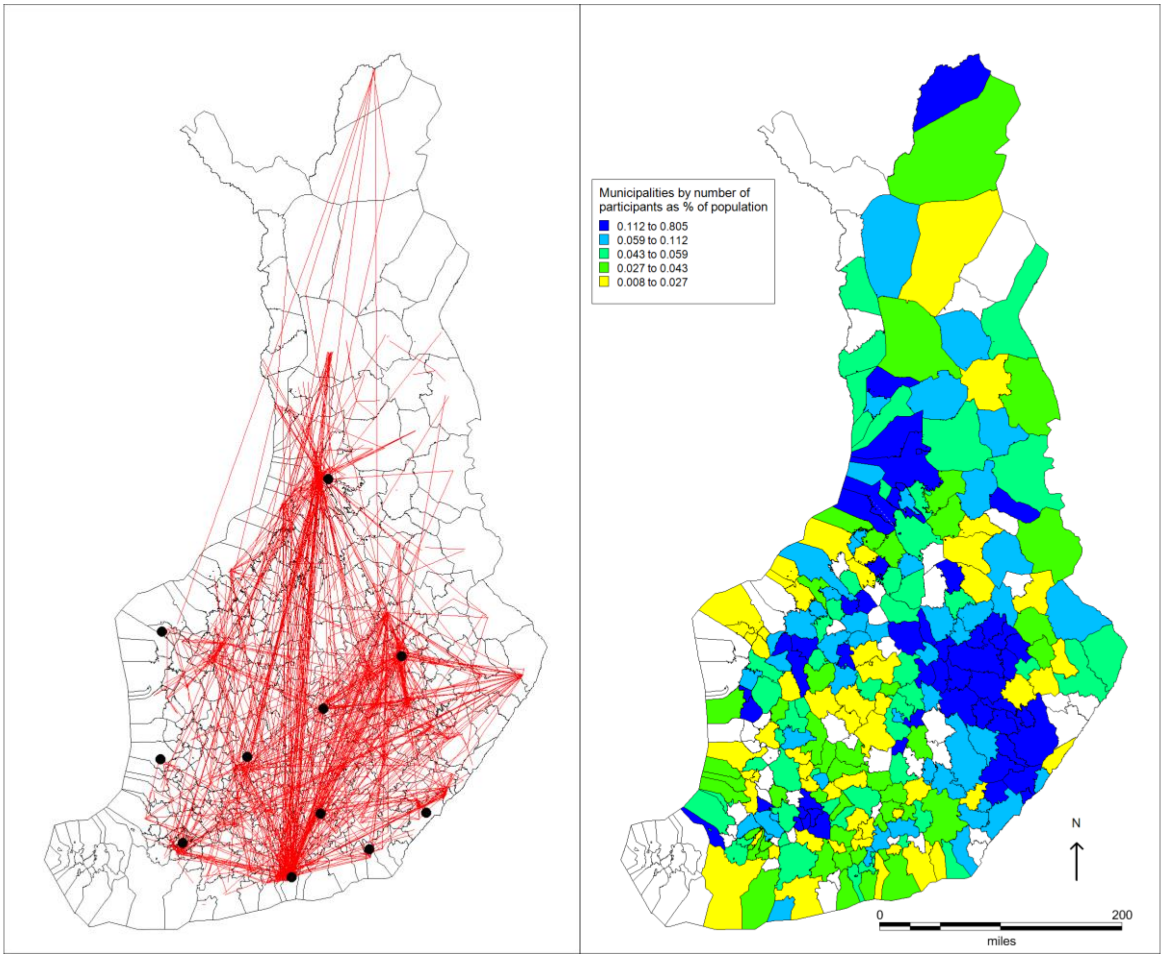



**Figure III**
**The Geographic Coverage of Wincapita**

This figure shows two maps illustrating the geographic coverage of Wincapita. Both maps are based on investors who were questioned by the police. The map on the left shows Wincapita's sponsoring relationships on a map of Finland. The red lines represent the sponsoring relationships and each line connects an investor and a sponsor. The investors' locations are based on the home addresses reported in the police interview documents. The lines are drawn based on all observations where both the investor's location and the sponsor's location are known. Investors whose location is outside Finland are excluded. The map also shows the borders of Finnish municipalities in the background. The black circles show the locations of the largest cities in Finland. The locations indicated by the circles include the Helsinki metropolitan area and the ten next largest cities. The map on the right shows statistics on the number of Wincapita investors as percentage of population in different municipalities. The percentages are calculated as the number of questioned investors in the municipality divided by the total population of the municipality. Different percentage range categories are denoted with different colors and the range corresponding to each color is reported in the figure. The municipalities with white background color do not have any investors in the sample. The true percentages of Wincapita investors in the population are higher than those reported in the map, because the reported percentages are based only on the investors who were questioned by the police.

**Full Sample** **Restricted Sample**

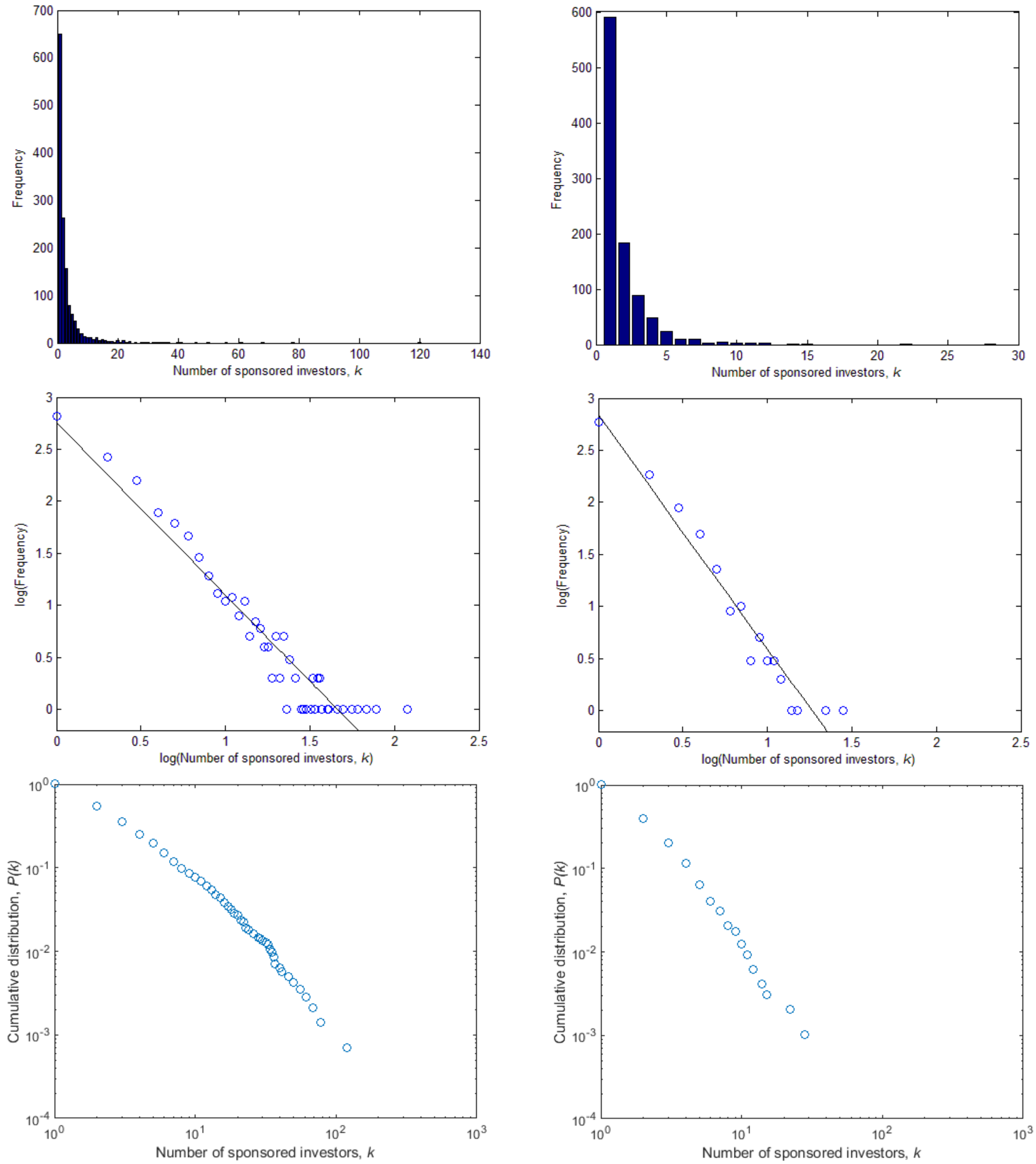



**Figure IV**
**Histograms and Log-Log Plots Based on the Number of Sponsored Investors**

These histograms and log-log plots show the distribution of the number of investors sponsored by a sponsor in the Wincapita sample. The graphs on the left are based on all observations in the sample and the graphs on the right are based on the restricted sample (investors who contacted the police by their own initiative). The histograms have frequency of observations on the y-axis and the number of sponsored investors on the x-axis. The log-log plots in the middle have log(frequency) based on the number of observations on the y-axis and log(number of sponsored investors) on the x-axis. They also include a fitted linear regression line based on the individual plots. The log-log plots in the bottom of the figure have the cumulative probability *P(k)* on the y-axis and the number of sponsored investors on the x-axis.

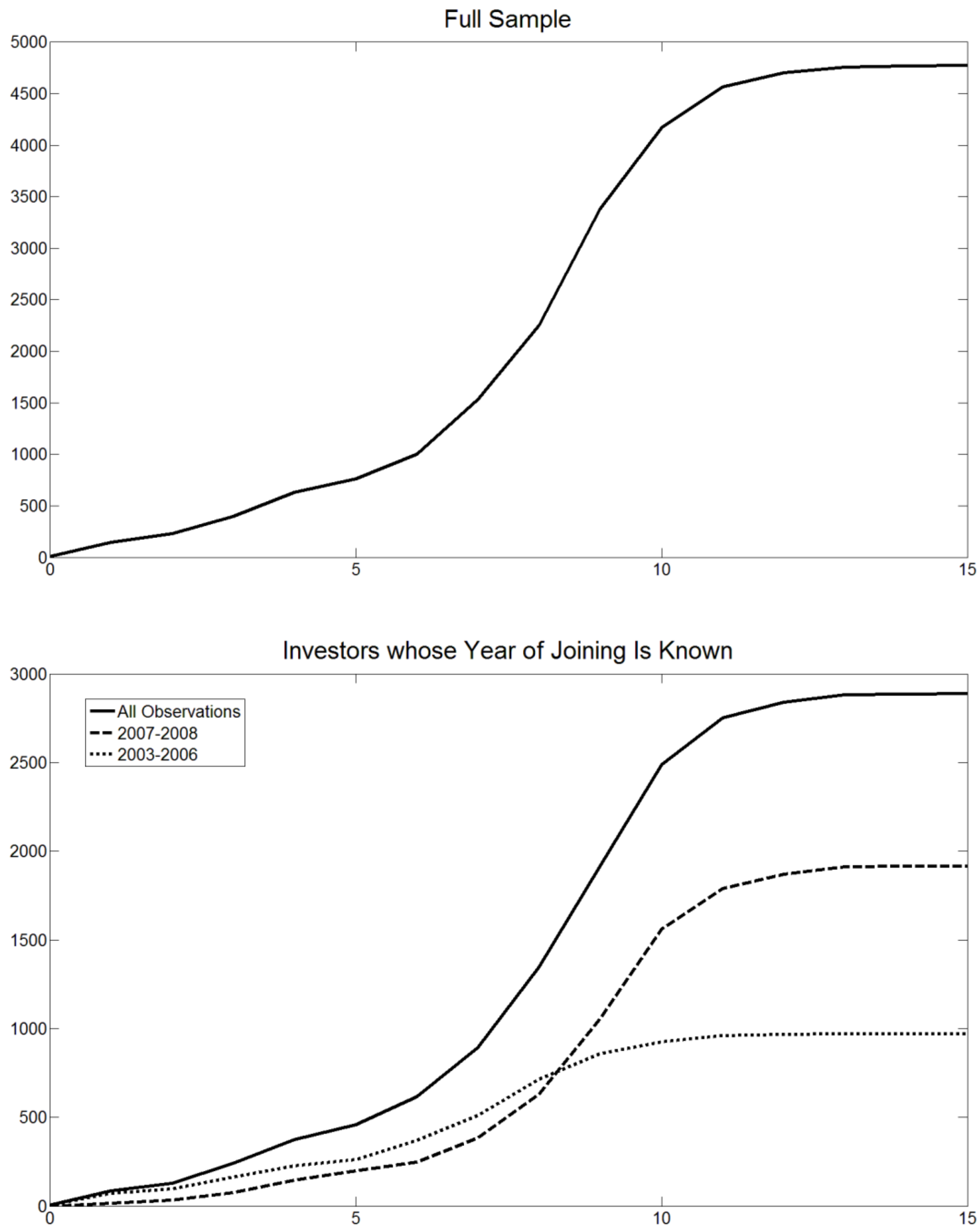



**Figure V**
**Network Distance and the Cumulative Number of Investors**

This figure shows graphs depicting the cumulative number of investors at different network distances. Network distance measures the number of sponsors between an investor and the starter of the Wincapita scheme (Hannu Kailajärvi) in the sponsoring network. The structure of the network is shown in Figure II. The graphs are based on investors for whom the chain of sponsors can be traced continuously to the originator of the scheme. The upper graph is drawn based on all sample investors and the lower graph is drawn based on investors whose year of joining is known. The lower graph includes separate curves based on subsamples of investors who joined Wincapita during 2003-2006 and 2007-2008.

**Table I**
**Statistics on Sample Investors and Data Availability**

This table reports statistics on the sample investors and data availability. Panel A provides statistics on data availability based on the number of observations with non-missing values for different data items. Place in network refers to investors whose sponsor is known or who have sponsored someone. Location is measured based on the person's street address. Panel B provides statistics on the personal characteristics of the investors.

Panel A: Data Availability (Number of Observations with Each Characteristic)

| Place in Network | Year of Joining | Total Invested Amount | First Invested Amount | Annual Income | Location | Education | All Seven Variables |
|---|---|---|---|---|---|---|---|
| 5,523 | 3,352 | 3,323 | 3,204 | 2,806 | 3,280 | 2,604 | 2,128 |

Panel B: Characteristics of the Investors

| | | | | | |
|---|---|---|---|---|---|
| Characteristics As % of Observations | Female | Entrepreneur | Age < 30 | Age > 60 | |
| | 19.72 % | 25.91 % | 9.52 % | 12.31 % | |
| Age at the End of 2007 | Average | 1st Quartile | Median | 3rd Quartile | |
| | 46.2 | 37 | 46 | 55 | |
| Education Level | Basic Education | Upper Secondary Education | Higher Education | Master's Degree | Doctoral Degree |
| | 12.8 % | 49.3 % | 22.4 % | 15.0 % | 0.5 % |
| Annual Income (Thousands of Euros) | Average | 1st Quartile | Median | 3rd Quartile | |
| | 47.6 | 23.9 | 33.7 | 48.2 | |

**Table II**

**Statistics on Wincapita Investments and Sponsoring Behavior**

This table reports statistics on Wincapita investors' invested and withdrawn amounts, year of joining, and the number of sponsored investors. Panel A provides statistics on the invested and withdrawn amounts based on the money transferred into and out of Wincapita in euros. Panel B reports the distribution of the sample investors' year of joining. Panel C reports statistics on the number of people sponsored by a sponsor (an investor who has invited at least one person into the scheme). Statistics in Panel C also include investors who were not personally questioned by the police, but who could be identified based on other investors' answers during the police investigation.

| Panel A: Invested and Withdrawn Amounts | Average | Min | 1st Quartile | Median | 3rd Quartile | Max |
|---|---|---|---|---|---|---|
| Total Invested Amount | 15,254 | 20 | 4,000 | 8,000 | 15,102 | 1,599,114 |
| First Invested Amount | 5,926 | 20 | 2,500 | 4,000 | 7,000 | 126,000 |
| Amount Withdrawn | 36,754 | 4 | 2,300 | 7,000 | 23,892 | 1,666,570 |
| Percentage of Investors Who Made Withdrawals | | | 26.5 | | | |
| Percentage of Investors Who Withdrew at Least €1,000 | | | 22.5 | | | |
| Percentage of Investors Who Withdrew at Least €10,000 | | | 10.6 | | | |

| Panel B: Year of Joining | 2003 | 2004 | 2005 | 2006 | 2007 | 2008 |
|---|---|---|---|---|---|---|
| Percentage of Investors | 1.4 | 7.0 | 10.0 | 12.0 | 43.2 | 26.3 |

| Panel C: Statistics on the Number of People Sponsored by a Sponsor | Average | 1st Quartile | Median | 3rd Quartile | Max | |
|---|---|---|---|---|---|---|
| | 3.7 | 1 | 2 | 3 | 120 | |
| Number of sponsored people | Only one | 2 to 4 | 5 to 10 | 11 to 20 | 21 to 40 | Over 40 |
| Observations | 650 | 498 | 179 | 64 | 26 | 8 |

**Table III**
**The Actual Wincapita Network Compared to a Simulated Random Network**

This table compare the actual Wincapita network to a simulated random network that has the same number of nodes. In the simulated network, the distribution of the number of sponsored people among sponsors follows a Poisson distribution. The mean of the Poisson distribution is equal to the corresponding mean in the actual Wincapita network. The simulated network is formed so that the first sponsor is the originator of the scheme and each newly sponsored investor is a sponsor himself with probability *p*, which is equal to the percentage of sponsors in the actual network. The number of new investors sponsored by each sponsor is drawn from the Poisson distribution. New nodes are added until the simulated network has the same number of nodes as the actual network. The only difference between the actual and the simulated networks is the shape of the distribution of the number of sponsored people. The statistics reported in the table include the maximum network distance required to reach the originator of the scheme, the number of investors within different network distance intervals relative to the originator of the scheme, and the average network distance to the originator of the scheme among all investors. The statistics for the simulated network are based on 10,000 simulation rounds, and mean values, percentiles, and standard deviations are reported for each statistic.

| | Actual Network | Simulated Network (Statistics Based on 1,000 Simulation Rounds) | | | | |
|---|---|---|---|---|---|---|
| | | Mean | St. Dev. | p25 | p50 | p75 |
| The Network Distance Required to Reach All Investors | 15 | 161.4 | 74.6 | 101.8 | 154 | 213 |
| The Number of Investors with a Network Distance 0-5 to the Originator | 764 | 57.6 | 37.6 | 32 | 46 | 71 |
| The Number of Investors with a Network Distance 0-10 to the Originator | 4,177 | 135.9 | 99.1 | 72 | 105 | 163 |
| The Number of Investors with a Network Distance 0-15 to the Originator | 4,774 | 235.3 | 185.7 | 121 | 173 | 180 |
| Average Network Distance to the Originator | 8.0 | 97.4 | 39.4 | 66.02 | 93.75 | 125.4 |

**Table IV**

**The Relationship between Network Distance and the Number of Sponsored Investors**

This table reports how the number of sponsored investors varies as a function of network distance. Network distance is calculated as the number sponsors between investor *i* and the originator of the scheme (Hannu Kailajärvi) in the sponsoring network. Panel A reports statistics based on fixed network distance intervals of four units and Panel B reports statistics based on network distance quartiles. The quartiles are formed by sorting all investors into groups based on their network distance. Reported statistics for each interval include the number of observations, the percentage of sponsors among the investors, the average number of sponsored people per investor, and the average number of sponsored people per sponsor.

Panel A: Network Distance and the Number of Sponsored Investors Based on Fixed Network Distance Intervals

| | | | | |
|---|---|---|---|---|
| Network Distance Range | 0-3 | 4-7 | 8-11 | 12-15 |
| Observations | 396 | 1,137 | 3,036 | 204 |
| Percentage of Sponsors | 0.250 | 0.287 | 0.237 | 0.123 |
| Average Number of Sponsored People per Investor | 1.576 | 1.425 | 0.809 | 0.319 |
| Average Number of Sponsored People per Sponsor | 6.303 | 4.969 | 3.417 | 2.600 |

Panel B: Network Distance and the Number of Sponsored Investors Based on Network Distance Quartiles

| | Quartile 1 | Quartile 2 | Quartile 3 | Quartile 4 |
|---|---|---|---|---|
| Network Distance Range | 0-6 | 7-8 | 9 | 10-15 |
| Observations | 1,001 | 1,250 | 1,132 | 1,390 |
| Percentage of Sponsors | 0.283 | 0.310 | 0.231 | 0.171 |
| Average Number of Sponsored People per Investor | 1.524 | 1.480 | 0.701 | 0.429 |
| Average Number of Sponsored People per Sponsor | 5.392 | 4.780 | 3.027 | 2.519 |

**Table V**
**The Personal Characteristics of Sponsors Compared to Non-Sponsors**

This table compares Wincapita sponsors to non-sponsors based on personal characteristics. The table also reports statistics for subsamples of sponsors who sponsored at least 10, 20, or 30 people. The reported statistics are average and median age (calculated at the end of 2007), average and median taxable income (in year 2007), the percentage of females, the percentage of people with higher than mandatory basic education, and the percentage of people with higher education (bachelor's degree or higher). The statistics are based on investors who were personally interviewed by the police. The table also reports simulation-based *p*-values for the difference between sponsors and non-sponsors. The two-sided *p*-value is based on a simulation where the same number of sponsors is selected randomly among all sample investors. It measures the probability that the difference between randomly assigned sponsors and non-sponsors is at least as large as the empirically observed difference based on 1,000 simulation rounds.

Panel A: Non-Sponsors Compared to Sponsors

| | Non-Sponsors | Sponsors | Simulation-Based *p*-Value for The Difference |
|---|---|---|---|
| Observations | 2,530 | 858 | |
| Median Age | 47 | 45 | 0.076 |
| Average Age | 46.4 | 45.0 | 0.005 |
| Median Income (Thousands of Euros) | 32.8 | 35.2 | 0.015 |
| Average Income (Thousands of Euros) | 45.7 | 53.1 | 0.017 |
| % Females | 21.3 | 12.4 | 0.000 |
| % with Higher Than Mandatory Education | 86.6 | 88.8 | 0.118 |
| % with Higher Education | 38.1 | 37.2 | 0.649 |

Panel B: Statistics for the Most Active Sponsors

| | Sponsors Who Sponsored at Least 10 People | Sponsors Who Sponsored at Least 20 People | Sponsors Who Sponsored at Least 30 People |
|---|---|---|---|
| Observations | 92 | 33 | 13 |
| Median Age | 48 | 48 | 56 |
| Average Age | 47.8 | 49.3 | 53.9 |
| Median Income (Thousands of Euros) | 52.0 | 116.5 | 127.4 |
| Average Income (Thousands of Euros) | 113.1 | 194.0 | 259.7 |
| % Females | 12.0 | 9.1 | 0.0 |
| % with Higher Than Mandatory Education | 81.3 | 74.1 | 75.0 |
| % with Higher Education | 33.3 | 25.9 | 33.3 |

**Table VI**

**Differences in Personal Characteristics within Individual Sponsor-Sponsored Investor Pairs**

This table compares investors to their sponsors. Comparisons are made based on age, education level, gender, geographic distance (in kilometers, based on the location of home), and income (taxable income in 2007). The education level has three categories: People with only mandatory basic education, people with upper secondary education, and people with higher education. The table reports statistics based on these variables, and for each statistic also the average value of the statistic based on a simulated network where investors are given a random place in Wincapita's network structure. The random placement is conducted without replacement and separately for each data item, so that only investors with non-missing values for the data item change places. The probabilities for observing the actual statistic in the simulation are also reported.

| | | | | | |
|---|---|---|---|---|---|
| Age Difference to Sponsor (Sponsor Age - Own Age) | | | | | |
| | Average | 1st Quartile | Median | 3rd Quartile | St. dev. |
| Sponsor Age - Own Age | 1.353 | -6 | 1 | 9 | 13.800 |
| Average Simulated Statistic | 0.048 | -11.927 | 0.049 | 12.016 | 17.244 |
| P(Simulated > Actual) | 0.092 | 0.000 | 0.091 | 0.993 | 1.000 |
| Education Difference to Sponsor | | | | | |
| Sponsor's Education Level | | Lower | Same | Higher | |
| (Percentage of Observation) | | 0.259 | 0.527 | 0.214 | |
| Average Simulated Statistic | | 0.297 | 0.403 | 0.299 | |
| Simulated Statistics Compared to Actual | | P(<Actual) | P(>Actual) | P(<Actual) | |
| | | 0.066 | 0.000 | 0.000 | |
| Gender Difference to Sponsor | | | | | |
| | | All Investors | Males | Females | |
| Percentage of Investors Sponsored by Females | | 0.087 | 0.061 | 0.189 | |
| Average Simulated Statistic | | 0.198 | 0.198 | 0.198 | |
| P(Simulated Statistic < Actual) | | 0.000 | 0.000 | 0.381 | |
| Geographic Distance to Sponsor | | | | | |
| | Average | 1st Quartile | Median | 3rd Quartile | St. dev. |
| Distance to Sponsor (Kilometers) | 76.401 | 4.420 | 16.058 | 80.459 | 130.4278 |
| Average Simulated Statistic | 350.285 | 211.777 | 360.120 | 479.118 | 184.1203 |
| P(Simulated Statistic < Actual) | 0.000 | 0.000 | 0.000 | 0.000 | 0.000 |
| Income Difference to Sponsor (Sponsor Income - Own Income) | | | | | |
| | Average | 1st Quartile | Median | 3rd Quartile | St. dev. |
| Sponsor Income - Own Income | 65.036 | -10.000 | 9.450 | 73.100 | 187.8186 |
| Average Simulated Statistic | 0.139 | -16.727 | 0.181 | 17.188 | 97.24908 |
| P(Simulated > Actual) | 0.000 | 0.000 | 0.000 | 0.000 | 0.007 |

**Table VII**
**Correlation in Personal Characteristics as a Function of Network Distance**

This table shows how correlation in personal characteristics within the Wincapita network varies as a function of social network distance. The correlation coefficients are calculated based on investor pairs where each investor *i* is linked to another investor *j* who joined the scheme before *i* and is part of the chain of sponsors that connects *i* to the originator of the scheme. Network distance is the number of sponsors between *i* and *j* in the network. E.g. the investor who has network distance zero to investor *i* is his sponsor, and the investor who has network distance one to investor *i* is his sponsor's sponsor. The table reports pairwise Pearson correlation coefficients and they are calculated based on all investor pairs with non-missing values for the data item. With continuous variables (age and income) I exclude the highest and lowest 1% of variable values in the sample to limit the effect of outliers. The variables measuring personal characteristics include Age (at the end of 2007), Education (a binary variable which takes the value one if the individual has higher than secondary education), Gender (a binary variable which takes the value one if the individual is female), and Income which measures taxable income for year 2007. *p*-values for each coefficient are reported below the coefficient value. ***, **, and * denote statistical significance at the 1%, 5%, and 10% level.

| Network Distance | 0 | 1 | 2 | 3 | 4 |
|---|---|---|---|---|---|
| **Age** | | | | | |
| Coefficient | 0.3287*** | 0.1580*** | -0.0147 | -0.0897*** | -0.0472* |
| *p*-value | 0.0000 | 0.0000 | 0.5495 | 0.0002 | 0.0677 |
| **Educaction** | | | | | |
| Coefficient | 0.2510*** | 0.1251*** | -0.0167 | -0.0207 | -0.0049 |
| *p*-value | 0.0000 | 0.0001 | 0.5746 | 0.4904 | 0.8860 |
| **Gender** | | | | | |
| Coefficient | 0.1898*** | 0.0519*** | 0.0158 | 0.0244 | -0.0142 |
| *p*-value | 0.0000 | 0.0037 | 0.4055 | 0.2111 | 0.4894 |
| **Income** | | | | | |
| Coefficient | 0.0600** | 0.0307 | -0.0453 | 0.0526* | -0.0742** |
| *p*-value | 0.0269 | 0.3380 | 0.1666 | 0.0981 | 0.0418 |