Well-to-do or Doing Well?
Empirical studies of wellbeing and development

ISBN: 978 90 316 0411 1

Cover design: Crasborn Graphic Designers bno, Valkenburg a.d. Geul

This book is no. 596 of the Tinbergen Institute Research Series, established through cooperation between Thela Thesis and the Tinbergen Institute. A list of books which already appeared in the series can be found in the back.

Well-to-do or Doing Well?
Empirical studies of wellbeing and development

Empirische studies over welzijn en ontwikkeling

Proefschrift

ter verkrijging van de graad van doctor aan de
Erasmus Universiteit Rotterdam
op gezag van de
rector magnificus

prof.dr. H.A.P. Pols

en volgens besluit van het College voor Promoties.

De openbare verdediging zal plaatsvinden op

vrijdag 19 december 2014 om 13:30 uur

door

Sanne Lise Blauw
geboren te Vlissingen

ERASMUS UNIVERSITEIT ROTTERDAM

Promotiecommissie

Promotor:	Prof.dr. P.H.B.F. Franses
Overige leden:	Prof.dr. R. Dur
	Prof.dr. T.J.S. Offerman
	Prof.dr. H.D. Webbink

Acknowledgments

In the past four years, I have lived in nine apartments in four different cities. I have travelled to sixteen countries. I have visited four continents. I have met many new people at conferences, summer schools, university visits, and, simply, on holidays. People who inspired me, taught me, broadened my view of the world. People who I still feel connected with, professionally and personally. People who made me feel at home, wherever I may have been.

My PhD journey started in Paris, where I worked at the OECD Development Centre. I am indebted to Andy, Johannes, Annalisa, Juan, and Chris for including me in their team and showing me the ins and outs of development economics. I am grateful to have met Roxana, who has become a great friend and who warmly welcomed me to her home in Bolivia. And, of course, I thank Eline for making my Parisian life colourful and adventurous.

After Paris, I went back to familiar territory: Erasmus University Rotterdam. Though I knew the gray buildings from my student days, I now saw them in a completely different light. I embarked on a new adventure: academia. I owe a lot to my supervisor Philip Hans, who - with his creative mind and support - helped me to discover the world of research. I thank Adriana and Emoke for introducing me to the art of teaching. Both at Erasmus and at the Tinbergen Institute, I have enjoyed the company of fellow PhD students: Eszter, Frank, Lerby, Lisette, Olivier, Saskia, Victor, and many others. A special mention goes to David, who showed me how fulfilling it can be to work together with a bright and friendly mind. Venturing outside the university walls, I met Hann, whom I have come to regard as a

mentor and a friend.

While meeting new people is exciting, I feel lucky to have stayed in touch with dear old friends. Loes, Lisa, and Laura, who keep my feet on Zeeland soil. The girls from JC Prik, who are always ready to welcome me back and have fun. Anne, who showed me that real friends think outside the box. Nathalie, with whom I feel at home even though we are not living together anymore. I cherish the long talks I have with Anna, and the idea that we're never done talking fills me with joy. Carlijn's enthusiasm, curiosity, and unconditional support make my life richer.

Whether at home or away, my family has always been close by. Over the years, the circle of relatives has grown. I have become attached to Gaby and Margalit, who make it every time a bit more difficult to say goodbye to Israel. I am fortunate to know Dominique, Marieke, and Steef, who have built new homes with the men of my family. The lightness and joy of Mies and Pia have added a new dimension to my life.

Some family members have been there from the very beginning. I will always remember my weekly trips to Beppe, whose wisdom, humbleness, and affection continue to be with me. Jurre - whether we are in English moors or Parisian bars, our Bolo home or smoky Taboe - your happiness, wit, and creativity always manage to lift me up. Hylke, I have always been proud to have you as my brother. As we grow older, we grow closer. I hope we will get many more wet feet together in the future. Mama, thank you for the tireless encouragement, open door, and never-ending conversations. You help me find my home in this world, which means everything. Papa, I think I could talk and walk with you into infinity. Let's never stop wondering.

Rei, you are the best thing that came out of doing a PhD. I could spend the rest of my life marvelling at the beautiful person you are. Your warmth, care, and respect enrich my life beyond words. Thank you for being there and staying there. I can't wait to see what life has in store for us.

Sanne Lise Blauw

Rotterdam, October 2014

Contents

Chapter 1

Introduction and Outline

"Not everything that counts can be counted, and not everything that can be counted counts."

– Sign in Einstein's office at Princeton University

1.1 Introduction

"How are you?" It is a question we hear on a daily basis. The answer seems straightforward: "Good." However, it is arguably one of the most complex questions someone can ask, as it requires the evaluation of every single facet of our lives: job, relationships, health, and so forth.

Many economists and policy makers are interested in quantitative measures of wellbeing. The concept of wellbeing is meant to represent the quality of life of an individual or group. In a way, it puts a number on the answer to "how are you?" And just like this question is difficult to answer, the measurement of wellbeing is difficult to do.

For a long time, economists have used monetary measures to proxy wellbeing, such as the gross domestic product (GDP). They are - at least in theory - relatively easy to calculate as the components

can be counted and added up. As the quotation at the start of this chapter alludes to, important aspects of life cannot necessarily be counted. Wellbeing is influenced by factors that are difficult or impossible to count, such as the quality of relationships or the way you feel when you wake up in the morning.

It is not obvious that people who are well-to-do are actually doing well. For this reason, non-monetary measures of wellbeing have become popular in economics over the past years. A good example is the young field of happiness economics, which has increasingly become accepted as valid research. The 2009 Stiglitz-Sen-Fitoussi report emphasised the importance to go "beyond GDP" to assess quality of life. This seminal report has sparked interest among policymakers. As a result, it is becoming more common for statistical agencies to collect data on happiness and other non-monetary dimensions of wellbeing.

In this thesis, I aim to investigate the drivers of wellbeing. Chapter 2 uses a traditional measure of wellbeing, household income, while subsequent chapters take a broader view of wellbeing, focusing on happiness and trust. Studying the determinants of wellbeing is difficult, even if there is agreement on the definition and measurement of the concept. First, the drivers of wellbeing may be influenced by wellbeing itself. For example, a high income might make someone happier, but it is also likely that happier people get jobs with higher salaries. Second, both wellbeing and its drivers may be affected by the same unobservable factors. For instance, an optimistic disposition is probably related to higher wellbeing and better health.

These two complications make it difficult to assess how certain factors impact wellbeing. In econometric jargon, statistical models will suffer from 'endogeneity' and estimated effects are biased. I face the issue of endogeneity in every chapter of this thesis and I deal with it in different ways. In Chapter 2 and 3, I estimate extra variables that, when included in the original model, get rid of potential endogeneity problems. In chapter 4, I analyse a group - a Dutch college sorority - with unique features that allow me to estimate effects in an unbiased manner. Chapter 5 copes with endogeneity most rigorously, employing a lab experiment in which participants are randomly assigned to a particular scenario.

This dissertation is based on data that I collected personally in Uganda, Bolivia, and The Netherlands. As an econometrician, I use quantitative data to investigate relevant research questions. However, conducting interviews allowed me to grasp the value of qualitative data. Beyond answering my questions, respondents often told me about the motivation for their answers or about their life in general. Such observations are important, especially for concepts that are difficult to quantify, such as wellbeing. I thus believe that wellbeing research would merit from a combination of quantitative and qualitative methods. Though I am not trained to perform rigorous qualitative research, I would like to give my respondents a voice beyond averages, coefficients, and standard errors. Therefore, most chapters include a quote from a respondent at the beginning.

1.2 Outline

Chapter 2 (based on joint work with Philip Hans Franses) analyses the effect of mobile phone use on household income in Uganda. Mobile phones have spread rapidly in developing countries over the last decade. It is now widely understood that cell phones have beneficial effects in these areas, but little is known about the impact at the household level.

A two-way causal relationship between economic development and telephone use is not unlikely, making it difficult to obtain unbiased estimates for the impact of mobile phones. After all, someone might become richer as a result of using a cell phone, but cell phone use also relies on having money. We use a novel econometric technique to handle the potential endogeneity in our statistical model. We model the correlation between the endogenous regressor and the error term with copulas, in line with Park and Gupta (2012). To the best of our knowledge, this is the first time that the copula method is applied in the economic development literature.

We find a positive impact of mobile phone use. More specifically, the proportion of mobile phone users in the household as well as the duration of mobile phone ownership matter for household income.

We find no evidence for a significant effect of the number of weekly mobile phone calls. Beyond conventional telephone use, we look at mobile banking and mobile search. Very few respondents use these services, but the descriptive statistics indicate that users are on average more economically developed.

In Chapter 3 I move beyond monetary wellbeing and look at life satisfaction (i.e. happiness, subjective well-being). Developing countries are underrepresented in the happiness literature, which is why I focus on Bolivia. In particular, I study the relationship between happiness and social reference groups. Social reference groups consist of individuals a person interacts with directly and frequently, such as neighbours and colleagues. It is known that individuals compare themselves to people around them and that their relative standing has an impact on their wellbeing. However, current studies rarely analyse social reference groups and rather focus on proxies of reference groups, such as people of the same age and gender.

In this chapter, I analyse self-reported, and thus subjective, comparisons to several reference groups. An obvious complication in studying the relation between two subjective variables - in my case subjective wellbeing and subjective comparisons - is the potential omitted variable bias resulting from a generally positive or negative disposition of respondents. For example, if an individual is naturally optimistic, she will answer more favourably to both subjective questions than a respondent with a more pessimistic outlook. My econometric methodology, similar to Mangyo and Park (2011), deals with this omitted variable bias by estimating a variable that measures disposition. I investigate and confirm the robustness of this method by using anchoring vignettes. The methodology does not deal with potential simultaneous causality of comparisons and life satisfaction. A two-way causal relationship is not unlikely, as happiness impacts factors that are important to relative standing, such as income. I can therefore not draw any conclusions about causality.

My findings support the importance of social reference groups and thereby confirm the relevance of these measures for research dealing with life satisfaction. More specifically, comparisons to family

members and (former) classmates are important determinants of subjective wellbeing. I find that other reference groups are associated with future life satisfaction, namely colleagues and neighbours. More generally, I find that the drivers of life satisfaction are in line with present research. The exception is marital status, for which I do not find a significant impact. In addition, I find that future satisfaction differs from present satisfaction, both in terms of distribution and determinants. Overall, I can conclude that both social reference groups and future life satisfaction provide additional valuable sources of information for subjective wellbeing research.

Chapter 4 focuses on yet another determinant of wellbeing: the happiness of the people around us. Social bonds tend to exist between individuals who are similar with respect to background, characteristics, and behaviour. Do similar people simply seek each other out (*homophily*) or do they also become more similar because of their interaction (*induction*)? This chapter aims to disentangle these two effects, using a unique data set from a Dutch college sorority. The contributions to current literature are twofold. First, it is one of the few network studies related to happiness and, to my knowledge, the first to study happiness in the setting of a fraternal organisation. Second, it is one of the few network studies on fraternal organisations and one of the first quantitative studies on Dutch fraternal organisations.

More specifically, I wish to study the effect of a peer's happiness on an individual's happiness. Statistical models are likely to suffer from endogeneity for several reasons. First, while the peer could affect the individual, it is also likely that the individual in turn influences the peer. Second, the happiness levels among individuals can be correlated for other reasons, such as a similar background or common changes to the environment (*confounding factors*). The setup of the sorority enables me to deal with these issues. Each year a new cohort joins the sorority and new members are expected to form groups within the cohort ('clubs'). Using data from before and after club formation, I am able to disentangle homophily and induction. Data on friendships allow me to control for confounding factors, assuming that friends are more likely to affect each other when they are closer socially.

Next to happiness, I look at other variables of interest, such as body mass index (BMI), alcohol use,

and study performance. I do not find evidence for homophily in happiness, behaviour or educational outcomes. However, parental income and existing networks are important in the sorting process: respondents tend to be connected if they come from the same city, both live in a sorority house or follow the same study major. Moreover, I find that happiness, alcohol use, and grades are subject to induction, whereas correlation in BMI and relationship status are due to confounding factors. Checks confirm the robustness of these effects. Since the induction effects are mainly driven by mutual friendships, they are possibly a consequence of homophily of friendships.

While the preceding two chapters focused on life satisfaction, Chapter 5 (based on joint work with David Smerdon) analyses another 'soft' element of wellbeing: trust. A large body of empirical research has found a negative correlation between trust and income inequality, and it has been suggested that the causal direction runs from inequality to trust. However, little is known about how this relationship depends on the *income distribution mechanism* in a society, which is likely to be relevant. For example, the impact of inequality is probably different in a society where citizens' incomes are based on merit than in a society where greed is rewarded.

Endogeneity is likely to occur in statistical models estimating the impact of inequality on trust. After all, inequality may affect trust, but trust may also determine inequality. Also, unobserved factors, such as cultural traits, could influence both trust and inequality. We therefore run a laboratory experiment in which we randomly assign participants to different scenarios. In our design, individuals are first placed in either a small, high-income class or a larger, low-income class, following a merit-based, greed-based or random allocation. A further treatment variable is the degree of inequality. Subjects then play a trust game against anonymous partners, including the elicitation of expectations with regards to the trustworthiness of their opponent.

Our main findings can be summarised as follows: Income inequality negatively impacts trust in a society in which income classes are determined randomly. When the income distribution mechanism is based on either merit or greed, however, we cannot conclude that changes in income inequality

affect trust within the group. Our findings are robust against selection effects. Also, we show that expectations, rather than sending behaviour, are correlated to survey measures of trust. We suggest that our results may be driven by the influence of the distribution mechanism on ingroup/outgroup effects.

Chapter 2

Off the hook: The Impact of Mobile Telephone Use on Economic Development in Uganda

Joint work with Philip Hans Franses

"Of course I have a mobile phone... Can I have your number?"

– John (70 years old)

2.1 Introduction

Mobile phone user rates have increased rapidly in the developing world. In Africa, for example, the mobile penetration rate was more than 65% in 2012, with an annual growth rate of 20% in the preceding five year (GSMA, 2013). Also, mobile services are becoming more advanced, now encompassing services such as mobile banking (m-banking). Mobile phone activity is likely to have a positive impact on economic development. Telephones enable market agents to collect price information more easily,

making markets more efficient (Eggleston et al., 2002). Also, mobile phones can help individuals to save time and reduce travelling costs (Mittal et al., 2009). M-banking can amplify these positive effects, by creating financial security, reducing the informal sector and encouraging entrepreneurship (Coyle, 2007).

Empirical evidence of the impact of mobile phone use can be split up in macro- and micro-economic studies (see Donner, 2008, for a comprehensive review). A precursor to the macro-economic research is Hardy (1980), who found a positive impact of fixed line telephones on economic growth. Waverman et al. (2005) find that this positive effect extends to mobile telecommunications and emphasise that this impact "may be twice as large in developing countries compared to developed countries" [p. 2]. Likewise, Kathuria et al. (2009) show that Indian states with higher mobile phone penetration have a higher economic growth.

Micro-economic impact studies tend to focus on the market effects of mobile phone use. Abraham (2007) shows that the use of mobile phones has a positive impact on economic development in the fishing industry in India. He finds that telephone use leads to increased market integration, gains in productivity, and reduction of price dispersion and price fluctuations. Similarly, Aker (2008) finds that the grain prices across Nigerian markets have been reduced by 20% as a result of mobile phone use.[1] Jensen (2007) demonstrates a "near-perfect adherence to the Law of One Price" [p. 879] in the South-Indian fisheries sector following the adoption of mobile phones. Muto and Yamano (2009) find that mobile phone coverage expansion increases market participation in Uganda for farmers growing perishable crops in remote areas.

In contrast to existing micro-economic research, our study focuses on the impact on individual households rather than on markets. A rare example of a household-level approach is Lee and Bellemare (2012), who study intra-household distribution of mobile phones in the Philippines. They find that farmers receive higher prices when the head of the household owns a mobile phone. It is not obvious

[1] Aker (2008) utilises the quasi-experimental character of cell-phone rollout and examines changing market power.

that market improvement has positive spillover effects on households that are not active market agents. Mobile phone use could even have negative effects on household income if phones are only used for purposes that are not income-generating nor cost-reducing.

We are also interested in additional mobile phone services, namely m-banking and 'mobile search.' M-banking allows customers to store, transfer and withdraw money using their mobile phones. Undoubtedly, M-PESA in Kenya is the most successful example of m-banking in a developing country (The Economist, 2009). Since M-PESA's launch in 2007, its customer base has grown to 23% of the Kenyan population and the transactions amount to 11% of Kenya's GDP (The Economist, 2010a). Most m-banking research is restricted to studying adoption (Laforet and Li, 2005) or use (Ivatury and Pickens, 2002; Porteous, 2007). Ivatury and Pickens (2002) and Porteous (2007) make the case that m-banking users in South Africa are more educated and richer than South Africans with a regular bank account. Ivatury and Mas (2008) address the 'early adopter' characteristics of users and predict that in the future m-banking will be used more by poor than by rich individuals. One of the rare quantitative impact studies is Mbiti and Weil (2011) who find that M-PESA[2] users have less informal savings and are more likely to use formal financial services.

Mobile search encompasses a range of Short Message Service (SMS) services that inform users *on request* of weather conditions, sports, news, agriculture, health and so on. Mobile search allows easy gathering of information, and it thus helps individuals to take more informed decisions. An example of such a service is Google SMS in Uganda, which enables users to ask for specific health (e.g. family planning) or agricultural information (e.g. tick control on livestock). The literature on mobile search is mainly of a technological nature, such as Jones et al. (2007) or focused on developed countries, like Kamvar and Baluja (2006).

An obvious problem in studying the effect of telephone use on development is the potential simultaneous causality. After all, economic development may influence the adoption and use of telephones.

[2]The 'M' in M-PESA refers to 'mobile' and pesa means 'money' in Swahili.

This simultaneous causal relationship implies endogeneity in the estimated regression, i.e. the standard error is correlated with the respective regressor. A common method to tackle this problem is instrumental variables (IV). IV relies upon variables that are both exogenous and strongly correlated with the endogenous variable. It is often difficult, if not impossible, to find valid instruments outside of an experimental setting. We therefore explore the use of a novel, instrument-free, method to deal with endogeneity. Following Park and Gupta (2012), we use copulas to model the correlation between the endogenous explanatory variable and the standard error. To the best of our knowledge, this is the first time this method is applied in economic development literature.

We apply the copula method to a unique cross-sectional data set. The first author conducted personal interviews with 196 heads of household in Uganda in 2010.[3] With a population of more than 32 million, Uganda had close to 10 million mobile subscriptions in March 2010 (Ugandan Communications Commission, 2010). The provider MTN was the first to introduce m-banking in Uganda in March 2009, soon to be followed by Zain in June 2009. MTN already reported its millionth m-banking subscriber in May 2010.[4] Several mobile search services are present in Uganda. Zain Uganda, for example, provides SMS information on request and for a charge on topics such as financial news, sports and horoscopes. Another example is Google SMS, which was launched in 2009 by MTN, Google and the Grameen Foundation. Google SMS offers, among others, free information on health and agriculture.[5]

We find a positive effect of the proportion of mobile phone users in the household as well as the duration of mobile phone ownership of the head of the household. We find no evidence that the number of weekly mobile phone calls has a significant impact on development. Very few respondents used m-

[3]Uganda is an interesting country for its pioneering role in telephones. It was the first to replicate the Village Phone program outside of Bangladesh, where it was initiated by the Grameen Foundation, and recently Uganda launched the Google SMS application. MTN villagePhone in Uganda is initiated by the telephone provider MTN Uganda and the Grameen Foundation. The program establishes so-called village phone operators (VPOs) in villages which lack electricity and provides them with a cell phone, which villagers can use for a small charge.

[4]See http://www.mtn.co.ug/About-MTN/News-Room/2010/May/MTN-MobileMoney-Customers-cros.aspx, last accessed on 27 August 2010.

[5]This information has been collected in cooperation with non-profit organisations.

banking or mobile search. The descriptive statistics, however, indicate that these services are associated with higher economic development at the household level.

The remainder of this paper is organised as follows. Section 2.2 discusses the data collection. Section 2.3.2 describes the econometric methodology and presents the empirical results. Finally, Section 2.4 provides conclusions, a discussion, and suggestions for future research.

2.2 Data collection

Our cross-sectional data set contains variables concerning economic development at the household level, telephone use and household characteristics. Unique data were collected by the first author in interviews with heads of households (N=196) on location in Uganda in March 2010.[6]

To account for the geographic diversity of Uganda, we selected three different areas as interview locations. Area one is located in the Central region of Uganda, in the Buikwe county in the Mukono district; area two is in Eastern Uganda, in the Tingey county in the Kapchorwa district; area three is in Western Uganda, in the Busongora county in the Kasese district. The map in Figure 2.1 indicates the locations of these areas. Interpreters assisted us in conducting the interviews in each of the three areas, translating to/from the languages of Luganda, Kuksabiny, and Logongo/Lutoro respectively. The remainder of this section describes the economic development, telephone use and household characteristics of our sample. Table 2.C2 in Appendix D presents descriptive statistics for urban areas and for our three sample areas separately.

[6]In case there were two heads of a household, for example when a family had a father and a mother, either would be accepted as an interviewee. Only one head of a household was interviewed.

2.2.1 Economic Development

It is difficult to measure economic development at the household level in developing countries, because clear income documentation is not always available. We therefore choose to use the Progress out of Poverty Index™ (PPI™) from (Chen2008). This index, which has been developed by the Grameen Foundation, is gaining popularity in microfinance as it provides a rough-and-ready measurement of development. The index is calculated using a scorecard containing a few simple questions, which allows for quick answering. Each item is associated with points, and the sum of the points for all questions is equal to the PPI™ for the household of the interviewee. As a formal scorecard is not yet available for Uganda, we use the PPI™ scorecard for the neighbouring country Kenya, as outlined in Chen et al. (2008). Table 3.2 displays this scorecard. In Appendix D, Table 2.C1 shows a statistical summary of our PPI™ data. The index ranges from 0 to 100, and a high (low) score indicates a low (high) probability that a household is below the poverty line. We interpret the PPI ™ as a measure of economic development, where we assume that development increases with index score.

Table 2.2 shows some key descriptive statistics of economic development in Uganda. For the full sample, the mean of 37.14 and the median of 37.00 of PPI™ differ only marginally, and the distribution thus seems symmetric. For areas one and two, however, the median is more than 4 points lower than the mean causing the distribution to be positively skewed, which is due to a few highly developed households. No household scores the full hundred points, implying that no single household is fully developed in terms of the PPI™. Both the lowest and the highest economically developed households are reported in area one, explaining the high standard deviation for this area. Figures 2.A1 to 2.A4 in Appendix B display the histograms of economic development for the full sample and for the three areas separately.

In addition to PPI™ scores, Table 2.2 shows the average poverty likelihood. This likelihood represents the probability (in percentages) that a household is below the poverty line.[7] The likelihoods for

[7]Chen et al. (2008) use the national poverty line, which is equal to the expenditures needed for food and non-food basic

the areas differ between 10% and 15% from the reported poverty levels, which can be explained by the fact that the poverty levels in Table 2.2 are given at the county level and not at the sub-county level.[8]

2.2.2 Telephone variables

The aim of our study is to measure the effect of mobile telephone use of the head of the household[9] on the economic development of his/her household. Telephone use is classified into 'basic telephone use' and 'advanced telephone use.' Public phone is assigned to the basic use category and incorporates both phone booths and pay phones, the latter being a telephone that can be used in exchange for a charge.[10] Basic telephone use also includes traditional use of mobile phone, which is both through voice (phone calls) and text (SMS). Advanced telephone use goes beyond traditional phone calls and SMS. Our study examines the applications of m-banking and mobile search, as described in the Introduction.

Table 2.3 shows the user rates of both basic and advanced telephone services in our sample. It is important to note that a 'mobile phone user' is defined as someone who *owns* a mobile phone. The mobile phone is more popular than the public phone, that is, 63% of the heads of households is a mobile phone user, while 43% is a public phone user Seventy-seven per cent of households has at least one mobile phone using member. On average, each household has one mobile phone user.

The low user rates of advanced services are striking, that is, only 12% for m-banking and only 5% for mobile search. These low percentages can perhaps be explained by the very recent introduction of both services. The currently small size of the samples of m-banking and mobile search users excludes advanced statistical analysis.

needs. They assign poverty likelihoods to each 5-point range of PPI™; the average of these likelihoods is presented in Table 2.2.

[8] For area one, the urban areas of the Buikwe county have a relatively low poverty level of 18%. However, the respondents in this area where from the town Nkokonjeru, which has poverty level 29%. In area three, the low poverty likelihood resulted from selecting respondents who, while living in rural areas, live close to the urban area of Kasese. Because detailed maps of area two are difficult to find, it is currently not possible to assign the correct poverty rates at sub county level.

[9] We use the variable the 'number of mobile phone users' to measure the telephone use of the entire household.

[10] Five percent of the individuals in the data set offered their personal mobile phone for public use for a charge.

Table 2.4 shows various descriptive statistics of variables measuring specifics of basic and advanced telephone use, such as frequency, purpose and expenditure. The statistics are reported for subsamples of users. Figures 2.A5 to 2.A9 in Appendix B display histograms of economic development for each subsample. Mobile phone users make a phone call almost every day, whereas public phone users make phone calls about two days a week. A mobile phone is more frequently used for phone calls than for SMS. Public phone and mobile phone, both voice and text, are mostly used to contact friends and family; less users have business reasons for their telephone use.

The histograms in Figures 2.B2 and 2.B3 in Appendix C show that distributions of ownership duration and airtime expenditure are positively skewed, explaining the high standard deviations relative to the mean and indicating potential outliers. The same applies to categorical variables like public phone call frequency, number of networks and SMS frequency (Figures 2.B1, 2.B4 and 2.B6 in Appendix C). Phone call frequency, however, is negatively skewed with many observations in the 'daily' category (Figure 2.B5).

Similar to public and mobile phone, m-banking is most often used to transfer to, or receive money from, friends and family members. On average, this service is used less than once a week and the amounts transferred range from 200 UGX[11] (0.07 EUR) to as much as 500,000 UGX (173.98 EUR). Mobile search services are used more than once a week and sports is the most popular information category.

2.2.3 Household characteristics

Table 2.5 presents sample characteristics of users of public phone, mobile phone, m-banking and mobile search.[12] Compared to the full sample, non-users - those respondents that do not use any telephone

[11] UGX is an acronym for the 'Ugandan Shilling.' The EUR/UGX exchange rate was equal to 2873.85 on 12 March 2010.

[12] Compared to national data, the sample of this study is different in several aspects. The household size exceeds the Ugandan average of 4.7 persons (Uganda Bureau of Statistics, 2006c), and the sample urbanisation rate is almost 4% higher than the rate of 12.3% for Uganda as a whole (Uganda Bureau of Statistics, 2006d). The average years of education is lower than the national average of 10 years (Central Intelligence Agency, 2009), while the literacy rate is more than 10% higher

service - are less economically developed, have a lower literacy rate and education level. Few of them live in urban areas, which in turn explains the high proportion of farmers. Non-users tend to be seven years older than users and most of them are female. Illiteracy, which is linked to a lower education level and a higher age,[13] appears to impede telephone use. Many of the findings regarding non-users also apply to public phone users, albeit to a lesser degree. Public phone users have a lower economic development, education level and literacy rate than the average of the total sample.[14]

Mobile phone users, on the other hand, score five points above the sample average of economic development. Mobile search and m-banking users score as much as eight points higher. Mobile search users' economic development shows more variation than that of other users.[15]

Almost all mobile phone users can read and write. The literacy rate of mobile search users is equal to 100%, which is not unexpected. All users, and m-banking and mobile search users in particular, have a higher education level than the sample mean. M-banking is more prevalent in rural areas, while mobile search is used more frequently in urban areas. The relatively high economic development, literacy and education level of advanced users correspond with the 'early adopter' characteristics that Ivatury and Mas (2008) mention in relation to m-banking users.

The average age is about the same across the different types of users, but the standard deviation is around two years lower for advanced users. Apparently, advanced services have not made their way (yet) to the youngest and the oldest categories of basic users.

The bottom panel of Table 2.5 refers to the feeling of well-being of the heads of households. The

than the population average of less than 69% (Uganda Bureau of Statistics, 2006b). This discrepancy is caused by a high literacy rate in area one (see Table 2.C2 in Appendix D). The proportion of farmers is lower than the national average of 71% (Uganda Bureau of Statistics, 2006a), caused by a high sample urbanisation rate.

[13]The average education level and age of illiterates is 3.4 years and 44.2 years, respectively, whereas these values for literates are 9.0 and 38.1.

[14]For public phone users who do not use a mobile phone, these results are more pronounced. With an education level of 6.51 years and literacy rate of 0.58, this category of users is only slightly outperforming the non-users. In terms of PPI™, which is 28.14 on average, these users are even doing worse than non-users.

[15]The high standard deviation is caused by the fact that the small mobile search sample reports PPI ™ values as low as 19 and as high as 86.

interviewees answered the following three questions on a five-point Likert scale (1=not at all, 5=extremely): "How important do you believe you are to others?"; "Do you believe you can accomplish in life what you want?"; and "Do you believe your children will succeed in education?" In general, interviewees gave fairly positive answers to all three questions. Public phone users are scoring below average with regards to well-being, and mobile phone users score above average. Mobile search users are scoring particularly high, rating around 0.5 points above the average for all three questions. Overall, the correlation between economic development and the average of three well-being variables is 0.24.

In sum, we see that mobile phone use, including advanced services, correlates positively with economic prosperity and with feelings of well-being. The sample of mobile search and m-banking users is small and we therefore cannot investigate this relationship further. In the next section we focus exclusively on the effects of basic telephone use.

2.3 Methodology and results

So far, we discussed the association between higher levels of economic development and telephone use. However, the direction of this relationship is uncertain. Indeed, more economic development may also facilitate the adoption of telephones. A two-way causal link implies endogeneity in regression estimation, i.e. the standard error is correlated with the endogenous regressor. If we wish to determine causality rather than correlation, our econometric methodology must take into account potential endogeneity.

2.3.1 Copula method

Basically, we consider the regression model

$$y = X\beta + \epsilon, \tag{2.1}$$

where y is a measure of economic development, X is an $n \times k$ matrix that includes an intercept, telephone variables, and control variables, β is a $k \times 1$ vector of coefficients and ϵ is a $n \times 1$ vector with errors. Simultaneity implies that regressors can be endogenous, which in econometric language means that

$$E(X'\epsilon) \neq 0, \tag{2.2}$$

where E is the expectations operator. Endogeneity leads to inconsistent and biased parameter estimates, which is a common problem in the literature on economic development, see Easterly and Levine (2002).

Normally, we would resort to Instrumental Variables (IV) estimation to obtain consistent parameter estimates. However, our dataset does not include variables that satisfy the exogeneity assumption of this method. Therefore, we turn to the "copula method" as proposed by Park and Gupta (2012), because this method does not require any instruments.

As stated before, endogeneity becomes an issue when a regressor and the structural error are correlated. However, if we can somehow model this correlation, we can account for it, and thereby deal with endogeneity. This is the key idea of the copula method, where the name originates from the use of copulas in order to estimate the joint density of the structural error and the endogenous regressor.

According to Sklar's theorem, every joint distribution can be written as a function of its margins and the other way around (see Theorem 1, Sklar, 1973). The 'copula' is the function that maps the - in our case two - cumulative distribution functions (CDFs) to their joint CDF. There are several copulas (Nelsen, 2006), but the most widely used is the Gaussian copula, which relies on the assumption that the variables have a joint normal distribution.

Assume that X consists of x_1 and X_2, which associate with one endogenous regressor and one or more exogenous regressors, respectively. We assume that the CDF of ϵ (F_ϵ) is a normal distribution with mean 0 and variance σ_ϵ^2.

Next, we use the Gaussian copula to get the joint CDF

$$G(x_1, \epsilon) = N(x_1^*, \epsilon^*), \tag{2.3}$$

where $x_1^* = \Phi^{-1}(F_x(x_1))$, $\epsilon^* = \Phi^{-1}(F_\epsilon(\epsilon))$, F_x signifies the CDF of x_1, Φ denotes the standard normal CDF and N is the bivariate standard normal distribution with correlation coefficient ρ. By differentiating Equation 2.3 we determine the joint probability density function

$$g(x_1, \epsilon) = \frac{\delta\delta G(x_1, \epsilon)}{\delta x_1 \delta \epsilon} f_x f_\epsilon, \tag{2.4}$$

where f_x and f_ϵ are the marginal densities of x_1 and ϵ respectively.

We could use this density to obtain the likelihood function and then consistently estimate the coefficient of the endogenous regressor using maximum likelihood estimation. However, we use an alternative estimation method by including x_1^* in Equation 3.1 and use OLS estimation (Park and Gupta, 2012). To see why this yields identical estimates, we write

$$\begin{pmatrix} x_1^* \\ \epsilon^* \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ \rho & \sqrt{1-\rho^2} \end{pmatrix} \begin{pmatrix} \nu_1 \\ \nu_2 \end{pmatrix}, \tag{2.5}$$

where ν_1 and ν_2 are independent random variables drawn from a standard normal distribution. We can only use the notation in Equation 2.5 if we assume that the joint distribution of x_1^* and ϵ^* is bivariate standard normal, which is the case when we use the Gaussian copula.

As the structural error in Equation 3.1 equals

$$\epsilon = F_\epsilon^{-1}(\Phi(\epsilon^*)) = \Phi_{\sigma_\epsilon^2}^{-1}(\Phi(\epsilon^*)) = \sigma_\epsilon \epsilon^*, \tag{2.6}$$

we can rewrite Equation 3.1 - using Equation 2.5 - as

$$y = x_1\beta_1 + X_2\beta_2 + \sigma_\epsilon(\rho x_1^* + (\sqrt{1-\rho^2})\nu_2), \tag{2.7}$$

Hence, by including x_1^* we can consistently estimate β_1, because we split the structural error from Equation 3.1 into two parts: (i) $\sigma_\epsilon \rho x_1^*$, which is correlated with x_1, but which we can estimate by including x_1^* as an additional regressor; and (ii) $\sigma_\epsilon(\sqrt{1-\rho^2})\nu_2$, which is uncorrelated with x_1.

This alternative estimation method can only be used when we assume normality of the structural error, and a bivariate normal distribution of this error and the endogenous regressor. This procedure resembles the Heckman error correction, as it also uses a generated regressor to deal with endogeneity bias (Heckman, 1978). As Park and Gupta (2012) show, this method can be generalised to J endogenous regressors x_j with $j = 1, ..., J$. We then include J generated regressors $x_j^* = \Phi^{-1}(F_x(x_j))$ in the OLS estimation.

The copula method cannot be used with endogenous regressors that are (i) binary or (ii) normally distributed. In both cases, the endogenous variable and its associated generated regressor will be highly correlated and standard errors will be high as a result of multicollinearity. In the binary case, there will even be perfect collinearity.

In practice, we estimate

$$y = x_1\beta_1 + X_2\beta_2 + \hat{x_1^*}\beta_3 + \eta, \tag{2.8}$$

where $\hat{x_1^*} = \Phi^{-1}(\hat{F}_x(x_1))$ and $\hat{F}_x$ is the CDF of x_1 as estimated by its frequency distribution.[16] Since $\hat{x_1^*}$ is a generated regressor, OLS estimation does not lead to the correct standard error for this

[16] We use that

$$\hat{F}_x(x_{1(l-1)}) < \Phi(x^*_{1(i \in S_l)}) < \hat{F}_x(x_{1l}), \tag{2.9}$$

where $i = 1, ..., N$ is an index for individuals, $l = 1, ..., L$ is an index for the ordered unique observations of the endogenous regressor, $\hat{F}_x$ is the empirical CDF of x_1, x_{1l} is the value of the endogenous regressor at index l and S_l is the

coefficient (Pagan, 1984). We therefore calculate the correct error with a bootstrap procedure. More specifically, we use a nonparametric bootstrap (Wooldridge, 2002) with 1,000 replications. We use Stata 11.2 to estimate the results; the Stata code and data are available upon request.

2.3.2 Results

Tables 2.6, 2.7, 2.8 and 2.9 display OLS and copula results for four endogenous regressors measuring quantity, duration and frequency of telephone use. The dependent variable is the logarithm of the PPI score as outlined in Section 2.2.1. We include years of education of the head of household, being a farmer, household size and area dummies as control variables.[17] Because we expect a concave relationship, we include demeaned squared terms for the telephone variables. We report heteroscedasticity robust standard errors.

Table 2.6 shows the *quantity effect* of mobile phones on household economic development. The proportion of mobile phone users has a significantly positive and concave effect on development, both in OLS and copula estimation. The average household[18] experiences an increase in development of 0.54% as a result of a 10% rise in users.[19] This effect is smaller than the effect of 0.63% for households that have no mobile phone users. For OLS, the effect estimated to be only 0.09% and 0.10% for the average and non-user household respectively.

The *duration effect* is estimated in Table 2.7. Both OLS and copula estimates imply a positive but

set of individuals where $x_i = x_l$. We calculate

$$\hat{x}^*_{1(i\in S_l)} = \Phi^{-1}(\hat{F}_x(x_{1(l-1)}) + (\hat{F}_x(x_1 l) + \hat{F}_x(x_{1(l-1)}))/2). \tag{2.10}$$

[17] All non-binary control variables - except for variables that measure proportions - are $log(1+x)$ transformed, thereby reducing skewness and heteroscedasticity. Individuals with missing values for the relevant variables included are deleted from the data set.

[18] In the average household, 24.4% of household members own a mobile phone.

[19] This effect is calculated as follows: [(linear effect)+(quadratic effect)*(mean proportion of mobile phone users)]*(% change in proportion of mobile phone users)$= (6.349 - 3.748 * 0.244) * 0.1 = 0.54$. The effect for a non-user household is simply the linear term multiplied by the % increase in users.

concave impact of the years of ownership on household development. At the mean of 2.5 years, an additional year of mobile phone use would increase development by 0.11%. Even though the squared term is significant, the magnitude is small compared to the coefficient of the linear term. This result indicates an almost linear relationship between duration and development. With an estimated impact of 0.05%, OLS again underestimates the effect.

Table 2.8 shows that the *frequency effect* disappears once we adequately deal with endogeneity. OLS estimation results in a significantly positive (linear) effect of the number of mobile phone calls, while copula estimation shows a coefficient that is not significantly different from zero. Looking at the high standard errors, we should be wary of a multicollinearity problem. For number of public phone calls we do not find significant coefficients for either method (Table 2.9).

2.4 Conclusion

We investigate the impact of telephone use on economic development in Uganda. Existing micro-economic studies focus on the market impacts of mobile telephone use. In contrast, we evaluate the impact *at the household level* and find that mobile phone use has a positive impact. We can therefore conclude that mobile phones have benefits extending beyond higher market efficiency. Policymakers should take these positive effects into account when deciding upon investments in mobile telecommunications.

More precisely, we find a positive impact of the proportion of mobile phone users in the household. Household members that own a mobile phone can receive information in a more efficient manner and can reduce travel costs. As a result, they contribute more to household development than they would have without a phone.

Also, the duration of mobile phone ownership of the head of the household significantly impacts household development. The explanation for this result is two-fold. First, it takes time to learn how

to use the mobile phone efficiently. For example, a shop owner needs to learn which persons to call to find out where to buy products at the lowest price. Second, the beneficial effects of mobile phone use may take time to take root and translate into enhanced development. Also, our index of economic development is inherently slow-moving, because its components relate to living standard and family circumstances that cannot all be changed quickly. For instance, one of the components of the index is the material of the walls of the house. Returning to our previous example, the shop owner will not build a new house overnight as a result of increased profits.

Last, we find no evidence of an impact of number of mobile phone calls on development. Extra phone calls may be purely for pleasure and do not contribute to development by increasing revenues or lowering costs. Public phone calls also do not increase development. For these phone calls, the (travelling) costs incurred to make the phone call might not offset the reaped benefits.

In general, we find rather small effects of telephone use. It takes time for development to be affected by mobile phone use and we expect the effects to be amplified over time. Users first need to learn and then need time to translate the benefits into actual development. It would therefore be helpful if future studies look at a longer time frame.

We also found that m-banking and mobile search use is positively correlated with economic development. The sample sizes were too small for further statistical analysis. We therefore suggest to collect more data for future research in this area.

For the first time in development literature, we use copulas to deal with the endogeneity caused by the potential simultaneous relation between telephone use and economic development. We found that OLS underestimated the effect of telephone use. Endogeneity concerns many topics in the field of development research. As valid instruments are often hard to find, instrument-free techniques offer a way forward for researchers investigating causality. The copula method used in this paper is such an instrument-free method. We believe that the application of copulas to other research areas merits further exploration.

Figure 2.1: The three geographical survey areas

Table 2.1: Economic development scorecard

Question	Answer	Points
1. How many household members are aged 25 or younger?	A: 3 or more	0
	B: 0, 1 or 2	8
2. How many household members aged 6 to 17 are currently attending school?	A: Not all	0
	B: All	8
	C: No children aged 6 to 17	21
3 What is the material of the walls of the house?	A: Mud/cow dung/grass/sticks	0
	B: Other	5
4. What kind of toilet facility does your household use?	A: Other	0
	B: Flush to sewer; flush to septic tank; pan/bucket; covered pit latrine; or ventilation improved pit latrine	2
5. Does the household own a TV?	A: No	0
	B: Yes	16
6. Does the household own a sofa?	A: No	0
	B: Yes	14
7. Does the household own a stove?	A: No	0
	B: Yes	12
8. Does the household own a radio?	A: No	0
	B: Yes	8
9. Does the household own a bicycle?	A: No	0
	B: Yes	5
10. How many head of cattle are owned by the household currently?	A: None or unknown	0
	B: 1 or more	9

Note: This table presents the scorecard used to calculate the Progress out of Poverty Index™ (PPI ™) of a household. The PPI™ is the sum of the scores of the answers to all ten questions in the table. The scorecard is a reproduction of the scorecard in Chen et al. (2008).

Table 2.2: Descriptive statistics of economic development for the full sample and for the three geographic areas

	Total	Area 1	Area 2	Area 3
Mean	37.14	40.38	31.69	40.92
Median	37.00	37.00	27.00	42.00
Maximum	86.00	86.00	67.00	67.00
Minimum	9.00	9.00	10.00	10.00
Std. Dev.	15.87	17.83	14.40	13.96
Observations	196	56	77	63
Average poverty likelihood (%)	35.98	35.62	44.61	35.13
County poverty level (%)	32.31	22.60	29.00	45.00

Note: This table presents descriptive statistics of economic development for the total sample and for the three separate geographic areas (See map in Figure 2.1). In addition, the table displays the average poverty likelihoods. Each household is assigned a poverty likelihood based on the PPI^{TM} score using results in Chen et al. (2008) and the average poverty likelihood is the mean of these values. The last row represents the county poverty levels as measured by Emwanu et al. (2007). The poverty level for area 1, which includes both rural and urban households, is a weighted average of the urban and rural rate. Areas 2 and 3 include only rural areas. Urban/rural areas are classified in accordance with Emwanu et al. (2007). Table 2.C2 in Appendix D shows the proportion of urban households for each area.

Table 2.3: Basic and advanced user rates

	Mean/ fraction	(Std. Dev.)
Basic use		
Public phone use (0=no, 1=yes)	0.43	
Mobile phone use, head of household (0=no, 1=yes)	0.63	
Mobile phone use, household (0=no, 1=yes)	0.77	
Users in household	1.35	(1.21)
Advanced use		
M-banking use (0=no, 1=yes)	0.12	
Mobile search use (0=no, 1=yes)	0.05	

Note: This table summarises user rates for the full sample giving the mean/fraction (and standard deviation in parentheses). All variables are measured at the head of the household level, except for one mobile phone use variable and the number of users in the household. Mobile phone use at the household level measures whether or not at least one member uses a mobile phone. A "mobile phone user" is defined as someone who *owns* a mobile phone.

Table 2.4: Descriptive statistics of basic and advanced telephone for subsamples of users

	Mean/ fraction	(Std. Dev.)
Basic use - public phone (85 obs)		
Phone call frequency (days a week)	1.67	(1.91)
Business phone calls (0=no, 1=yes)	0.30	
Friends/family phone calls (0=no, 1=yes)	0.90	
Basic use - mobile phone (125 obs)		
Ownership duration (years)	4.10	(3.53)
Airtime expenditure (10000 UGX per month)	2.21	(2.66)
Number of networks used	1.26	(0.49)
Phone call frequency (days a week)	6.09	(1.80)
SMS frequency (days a week)	1.51	(2.11)
Business phone calls (0=no, 1=yes)	0.59	
Friends/family phone calls (0=no, 1=yes)	0.90	
Business SMS (0=no, 1=yes)	0.25	
Friends/family SMS (0=no, 1=yes)	0.50	
Advanced use - m-banking (24 obs)		
Frequency (days a week)	0.70	(0.54)
Average transaction	8.29	(10.14)
Business transactions (0=no, 1=yes)	0.25	
Friends/family transactions (0=no, 1=yes)	0.63	
Advanced use - mobile search (10 obs)		
Frequency (Days a week)	2.46	(2.88)
Agriculture information (0=no, 1=yes)	0.10	
Health information (0=no, 1=yes)	0.10	
Sports information (0=no, 1=yes)	0.50	
Trade information (0=no, 1=yes)	0.10	
Weather information (0=no, 1=yes)	0.00	

Note: This table shows the variables used to measure mobile phone, public phone, m-banking and mobile search use. The table presents a variable's mean or fraction (and standard deviation in parentheses) for the subsample of users of the relevant service. All variables are measured at the level of the head of the household. The acronym 'UGX' denotes Ugandan Shilling. The EUR/UGX exchange rate was equal to 2873.85 on 12 March 2010. A "mobile phone user" is defined as someone who *owns* a mobile phone.

Table 2.5: Sample and sub sample descriptive variables

	Total	Non-users	Public phone users	Mobile phone users	M-banking users	Mobile search users
Observations	197	29	85	125	24	10
Economic development	37.14 (15.87)	29.07 (11.50)	33.26 (14.90)	42.15 (15.80)	42.25 (14.32)	45.40 (22.17)
Household variables						
Household size	6.72 (2.92)	6.14 (2.66)	6.54 (2.99)	6.81 (3.12)	5.91 (2.83)	5.75 (2.55)
Urban area (0=no, 1=yes)	0.16	0.03	0.14	0.19	0.09	0.63
Proportion males in household	0.49	0.45	0.50	0.49	0.49	0.53
Head of household variables						
Age in years	38.77 (12.21)	44.14 (19.56)	37.57 (10.50)	37.93 (9.85)	37.09 (8.12)	37.25 (7.74)
Years of education	7.87 (3.84)	5.03 (4.02)	8.15 (3.77)	8.96 (3.49)	9.22 (3.55)	10.88 (2.17)
Gender (0=female, 1=male)	0.48	0.17	0.56	0.53	0.57	0.63
Farmer (0=no, 1=yes)	0.44	0.72	0.51	0.27	0.26	0.00
Salesperson (0=no, 1=yes)	0.20	0.03	0.17	0.24	0.17	0.50
Marital status (0=not married, 1=married)	0.80	0.79	0.81	0.82	0.87	0.63
Literacy (0=illiterate, 1=literacy)	0.80	0.45	0.77	0.95	0.96	1.00
Well-being variables						
Importance (1=not at all, 5=extremely)	4.26 (0.72)	4.24 (0.91)	4.21 (0.76)	4.29 (0.67)	4.02 (0.41)	4.88 (0.35)
Accomplishment (1=not at all, 5=extremely)	3.89 (1.06)	3.41 (1.39)	3.72 (1.06)	4.12 (0.87)	3.80 (0.58)	4.50 (1.07)
Children (1=not at all, 5=extremely)	4.22 (0.90)	3.68 (1.16)	4.16 (0.87)	4.46 (0.67)	4.30 (0.93)	4.75 (0.46)

Note: This table presents sample characteristics on well-being for the full sample and for the subsamples of non-users and users of public phone, mobile phone, m-banking and mobile search. "Non-users" are those individuals who use neither public nor mobile phone. The means or fractions of survey questions are reported with their standard deviations in parentheses. Economic development is measured in terms of PPI^{TM} as is explained earlier. The urban/rural classification of Emwanu et al. (2007) is used to determine the proportion of households in urban areas. To measure well-being, three questions were asked to which the interviewees had to give an answer on a scale from 1 ("not at all") to 5 ("extremely"). These questions were: "How important do you believe you are to others?" ('Importance'); "Do you believe you can accomplish in life what you want?" ('Accomplishment'); and "Do you believe your children will succeed in education?" ('Children'). A "mobile phone user" is defined as someone who *owns* a mobile phone.

Table 2.6: OLS and copula results for economic development with endogenous regressor proportion of mobile phone users in household

	ols	copula
Constant	3.168***	2.003***
	(0.18)	(0.69)
Proportion mobile phone users in household	1.016***	6.349**
	(0.19)	(3.11)
Proportion mobile phone users in household - squared	-0.545*	-3.748**
	(0.31)	(1.89)
Education, head of household (years)	0.090*	0.097**
	(0.05)	(0.05)
Farmer	-0.119*	-0.110
	(0.06)	(0.07)
Household size	0.043	0.075
	(0.06)	(0.06)
Area 1	-0.020	-0.015
	(0.07)	(0.08)
Area 2	-0.172**	-0.180**
	(0.07)	(0.07)
Generated regressor 1		-1.137*
		(0.67)
Generated regressor 2		-0.043
		(0.05)
Observations	193	193
R-squared	0.361	0.372

Note: 'ols' refers to OLS estimation; 'copula' refers to estimation using the copula method as outlined in Park and Gupta (2012). We report heteroscedasticity robust standard errors in parentheses. *,**,*** indicate significance at the 10%, 5% and 1% level respectively. In the copula estimation we include 'Generated regressor 1' and 'Generated regressor 2' corresponding to the generated regressors for the endogenous variable and its demeaned squared term respectively. The standard errors of the generated regressors are calculated with a nonparametric bootstrap procedure.

Table 2.7: OLS and copula results for economic development with endogenous regressor years of mobile phone ownership of head of household

	ols	copula
Constant	3.589***	3.444***
	(0.16)	(0.22)
Years of mobile phone ownership	0.067***	0.117*
	(0.01)	(0.06)
Years of mobile phone ownership - squared	-0.005***	-0.004*
	(0.00)	(0.00)
Education, head of household (years)	0.089*	0.090*
	(0.05)	(0.05)
Farmer	-0.171***	-0.181***
	(0.06)	(0.07)
Household size	-0.132**	-0.123**
	(0.05)	(0.05)
Area 1	-0.016	-0.015
	(0.08)	(0.08)
Area 2	-0.132*	-0.126
	(0.07)	(0.08)
Generated regressor 1		-0.176
		(0.20)
Generated regressor 2		-0.086*
		(0.05)
Observations	192	192
R-squared	0.326	0.334

Note: 'ols' refers to OLS estimation; 'copula' refers to estimation using the copula method as outlined in Park and Gupta (2012). We report heteroscedasticity robust standard errors in parentheses. *,**,*** indicate significance at the 10%, 5% and 1% level respectively. In the copula estimation we include 'Generated regressor 1' and 'Generated regressor 2' corresponding to the generated regressors for the endogenous variable and its demeaned squared term respectively. The standard errors of the generated regressors are calculated with a nonparametric bootstrap procedure.

Table 2.8: OLS and copula results for economic development with endogenous regressor mobile phone calls per week by head of household

	ols	copula
Constant	3.441***	1.598
	(0.20)	(3.43)
Mobile phone calls per week	0.043***	0.047
	(0.01)	(0.26)
Mobile phone calls per week - squared	0.009	0.178
	(0.01)	(0.23)
Education, head of household (years)	0.103*	0.093*
	(0.05)	(0.05)
Farmer	-0.184***	-0.184***
	(0.07)	(0.07)
Household size	-0.136**	-0.126**
	(0.05)	(0.05)
Area 1	-0.058	-0.053
	(0.08)	(0.08)
Area 2	-0.135*	-0.130
	(0.08)	(0.08)
Generated regressor 1		-0.132
		(1.51)
Generated regressor 2		-0.901
		(1.39)
Observations	193	193
R-squared	0.292	0.306

Note: 'ols' refers to OLS estimation; 'copula' refers to estimation using the copula method as outlined in Park and Gupta (2012). We report heteroscedasticity robust standard errors in parentheses. *,**,*** indicate significance at the 10%, 5% and 1% level respectively. In the copula estimation we include 'Generated regressor 1' and 'Generated regressor 2' corresponding to the generated regressors for the endogenous variable and its demeaned squared term respectively. The standard errors of the generated regressors are calculated with a nonparametric bootstrap procedure.

Table 2.9: OLS and copula results for economic development with endogenous regressor public phone calls per week by head of household

	ols	copula
Constant	3.661***	3.568***
	(0.17)	(0.20)
Public phone calls per week	-0.026	0.175
	(0.05)	(0.19)
Public phone calls per week - squared	0.001	-0.021
	(0.01)	(0.02)
Education, head of household (years)	0.151***	0.156***
	(0.06)	(0.06)
Farmer	-0.241***	-0.233***
	(0.07)	(0.06)
Household size	-0.114**	-0.119**
	(0.06)	(0.05)
Area 1	-0.103	-0.096
	(0.08)	(0.08)
Area 2	-0.240***	-0.225***
	(0.07)	(0.07)
Generated regressor 1		-0.236
		(0.17)
Generated regressor 2		-0.010
		(0.09)
Observations	193	193
R-squared	0.249	0.271

Note: 'ols' refers to OLS estimation; 'copula' refers to estimation using the copula method as outlined in Park and Gupta (2012). We report heteroscedasticity robust standard errors in parentheses. *,**,*** indicate significance at the 10%, 5% and 1% level respectively. In the copula estimation we include 'Generated regressor 1' and 'Generated regressor 2' corresponding to the generated regressors for the endogenous variable and its demeaned squared term respectively. The standard errors of the generated regressors are calculated with a nonparametric bootstrap procedure.

2.A Histograms - economic development

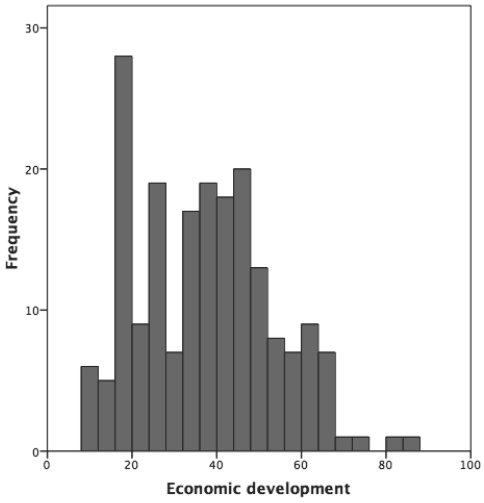


Figure 2.A1: Histogram of economic development, as measured by PPI [TM], for the full sample.

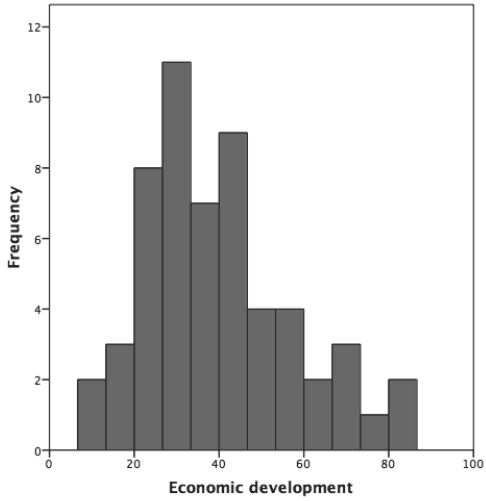


Figure 2.A2: Histogram of economic development, as measured by PPI ^{TM}, for area 1.

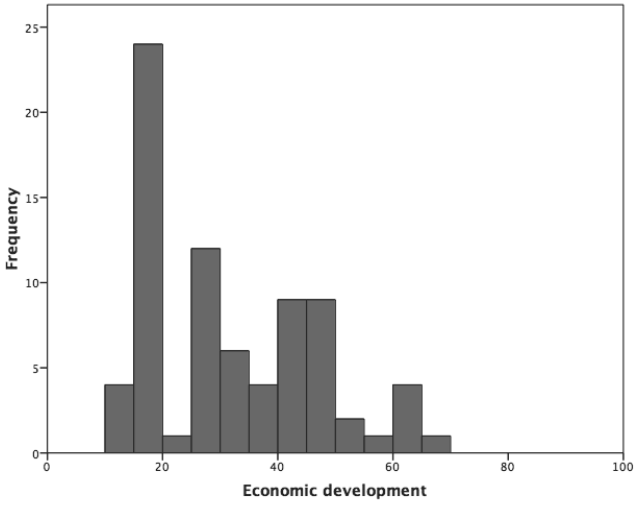


Figure 2.A3: Histogram of economic development, as measured by PPI ^{TM}, for area 2.

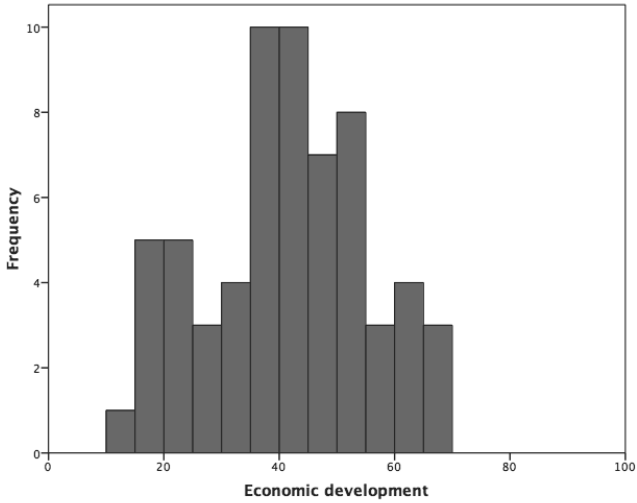


Figure 2.A4: Histogram of economic development, as measured by PPI ^{TM}, for area 3.

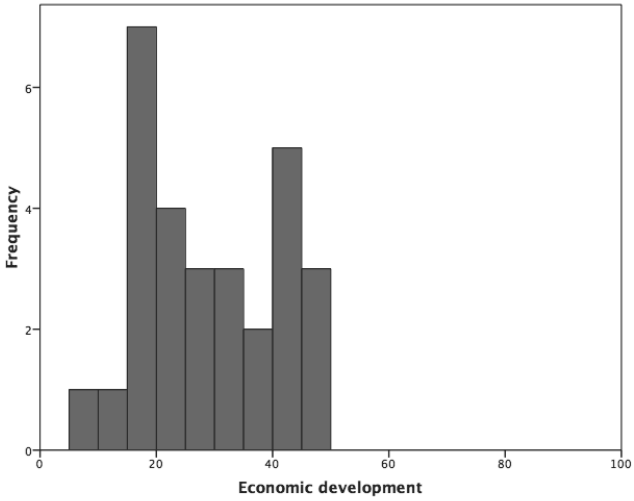


Figure 2.A5: Histogram of economic development, as measured by PPI ^{TM}, for non-users.

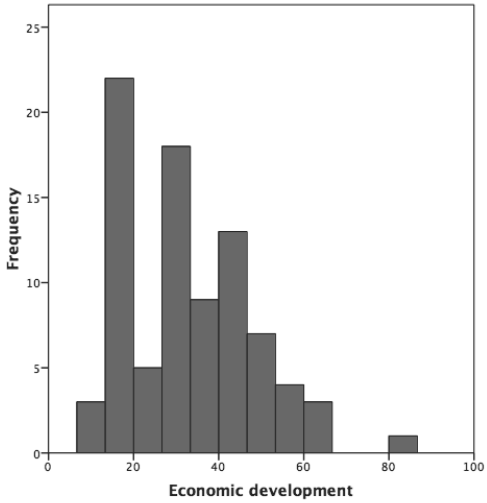


Figure 2.A6: Histogram of economic development, as measured by PPI ^{TM}, for public phone users.

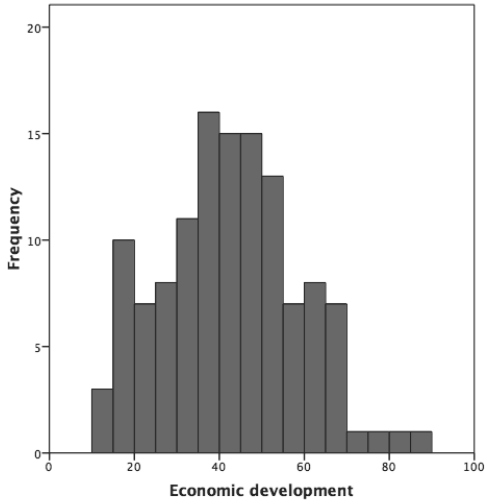


Figure 2.A7: Histogram of economic development, as measured by PPI ^{TM}, for mobile phone users.

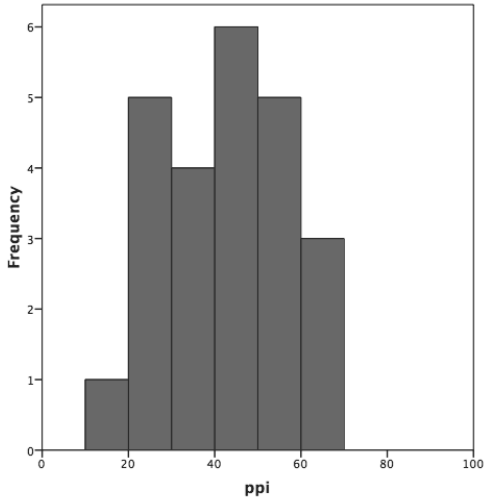


Figure 2.A8: Histogram of economic development, as measured by PPI ™, for m-banking users.

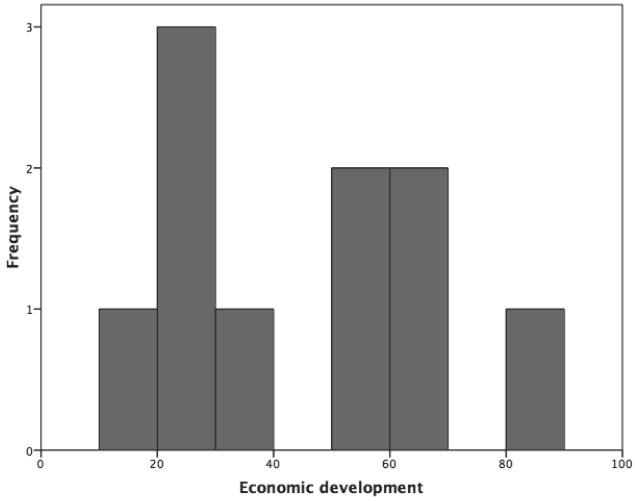


Figure 2.A9: Histogram of economic development, as measured by PPI ^{TM}, for mobile search users.

2.B Histograms - telephone use

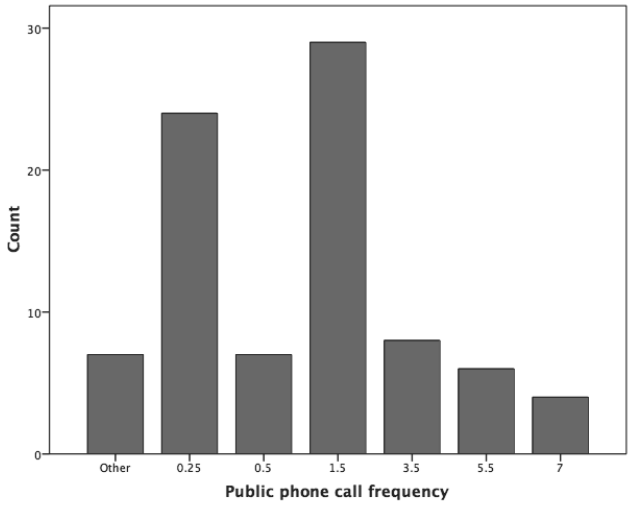


Figure 2.B1: Histogram of public phone call frequency for the subset of public phone users

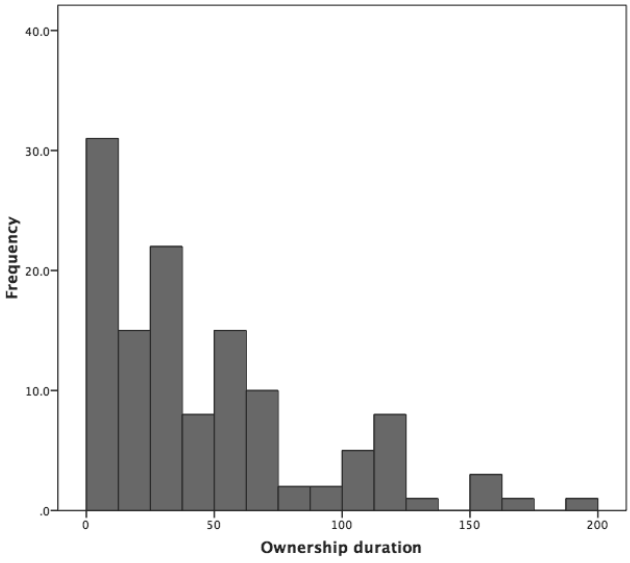


Figure 2.B2: Histogram of ownership duration for the subset of mobile phone users

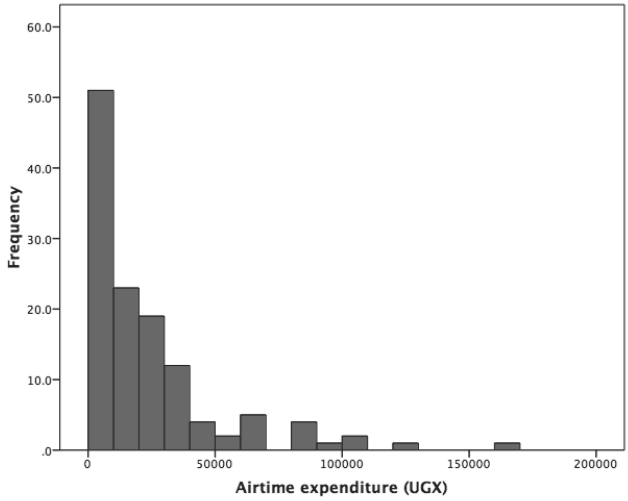


Figure 2.B3: Histogram of airtime expenditure for the subset of mobile phone users

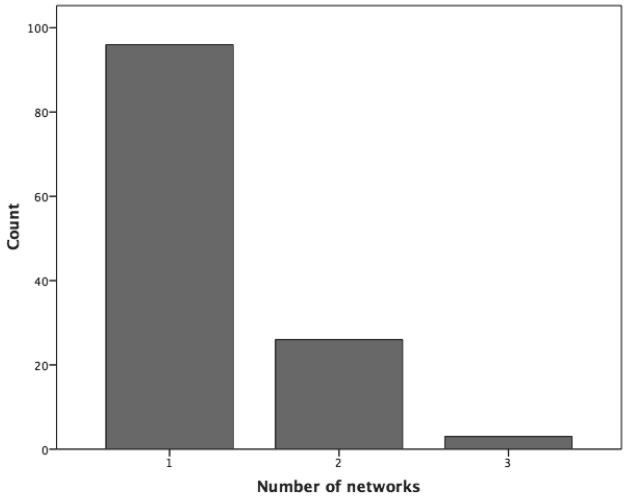


Figure 2.B4: Histogram of number of networks used for the subset of mobile phone users

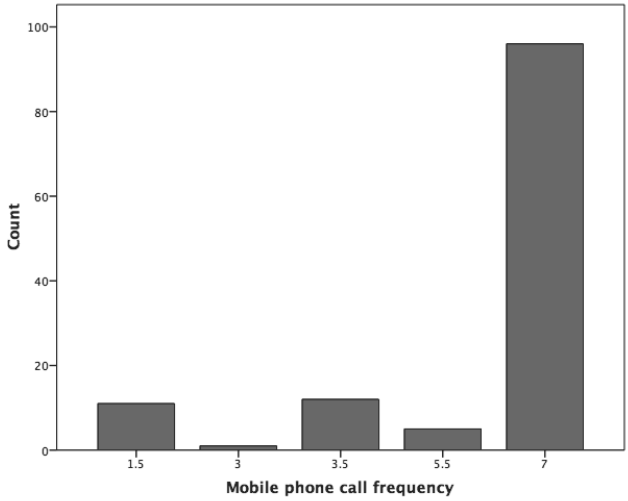


Figure 2.B5: Histogram of mobile phone call frequency for the subset of mobile phone users

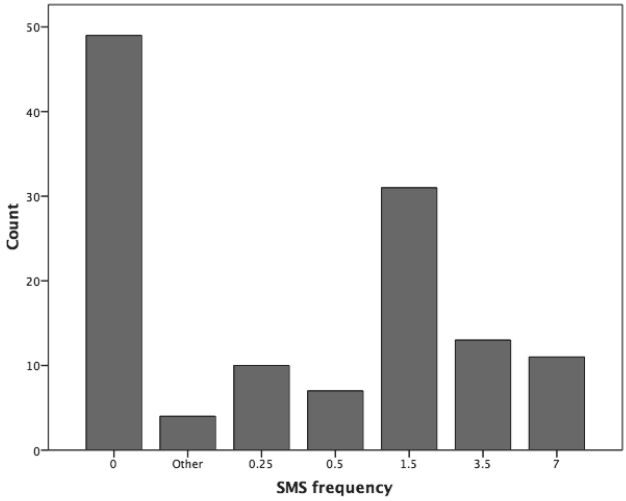


Figure 2.B6: Histogram of SMS frequency for the subset of mobile phone users

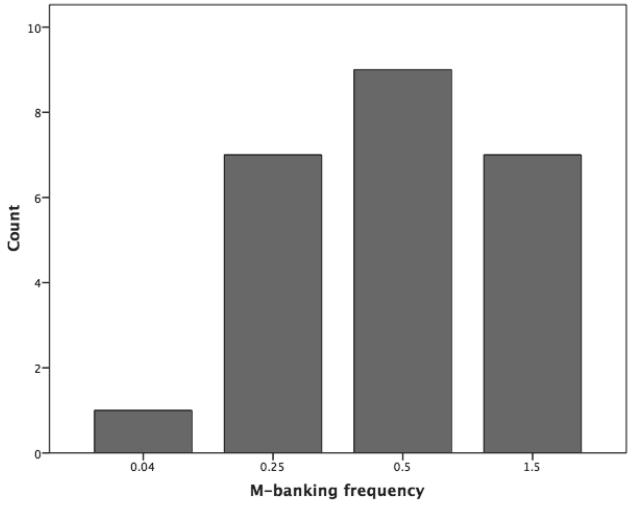


Figure 2.B7: Histogram of m-banking frequency for the subset of m-banking users

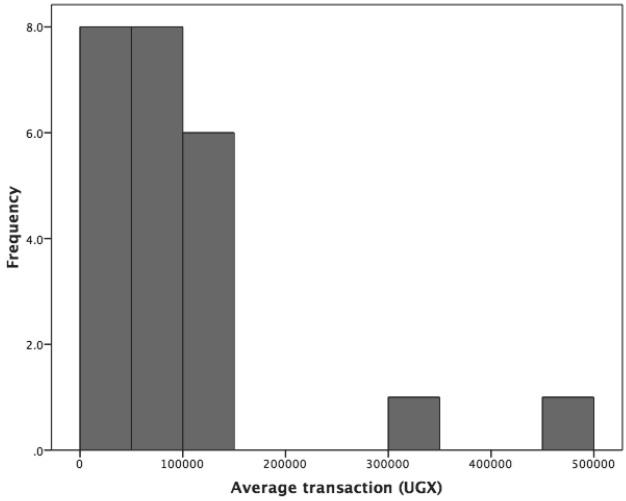


Figure 2.B8: Histogram of average transaction for the subset of m-banking users

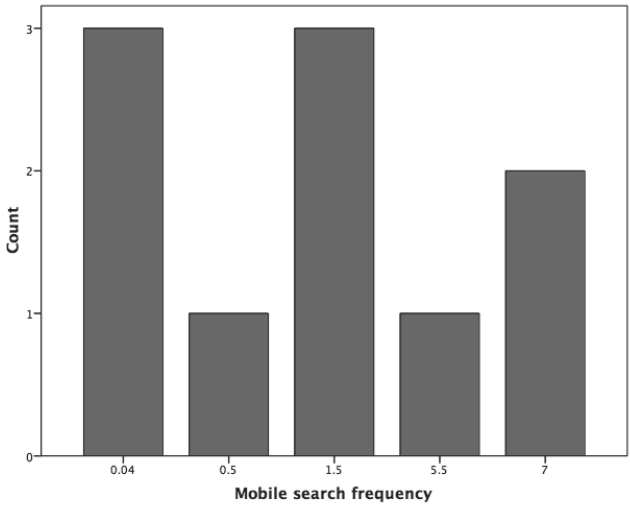


Figure 2.B9: Histogram of mobile search frequency for the subset of mobile search users

2.C Descriptive statistics for geographic areas

Table 2.C1: Frequency of scorecard answers for the full sample and the geographic areas

Question	Answer	Total	Area 1	Area 2	Area 3
1. How many household members are aged 25 or younger?	A: 3 or more	0.79	0.61	0.84	0.87
	B: 0, 1 or 2	0.21	0.39	0.16	0.13
2. How many household members aged 6 to 17 are currently attending school?	A: Not all	0.04	0.04	0.03	0.06
	B: All	0.77	0.79	0.73	0.79
	C: No children aged 6 to 17 and 21	0.19	0.18	0.25	0.14
3 What is the material of the walls of the house?	A: Mud/cow dung/grass/sticks	0.46	0.09	0.95	0.19
	B: Other	0.54	0.91	0.05	0.81
4. What kind of toilet facility does your household use?	A: Other	0.04	0.11	0.00	0.02
	B: Flush to sewer; flush to septic tank; pan/bucket; covered pit latrine; or ventilation improved pit latrine	0.96	0.89	1.00	0.98
5. Does the household own a TV?	A: No	0.87	0.80	0.94	0.86
	B: Yes	0.13	0.20	0.07	0.14
6. Does the household own a sofa?	A: No	0.60	0.63	0.73	0.43
	B: Yes	0.40	0.38	0.27	0.57
7. Does the household own a stove?	A: No	0.98	0.96	1.00	0.97
	B: Yes	0.02	0.04	0.00	0.03
8. Does the household own a radio?	A: No	0.25	0.29	0.33	0.13
	B: Yes	0.75	0.71	0.68	0.87
9. Does the household own a bicycle?	A: No	0.75	0.73	0.95	0.51
	B: Yes	0.26	0.27	0.05	0.49
10. How many head of cattle are owned by the household currently?	A: None or unknown	0.39	0.45	0.26	0.51
	B: 1 or more	0.61	0.55	0.74	0.49

Note: This table presents the frequency of answers given to each question on the PPI ^{TM} scorecard. The scorecard in this table is a reproduction of the scorecard in Chen et al. (2008).

Table 2.C2: Sample descriptive variables for urban areas and the three geographic areas

	Urban	Area 1	Area 2	Area 3
Observations	31	57	77	63
Household variables				
Household size	5.31 (2.05)	5.98 (3.20)	6.86 (2.74)	7.16 (2.82)
Urban area (0=no, 1=yes)	1.00 (NA)	0.54 (0.50)	0.00 (NA)	0.00 (NA)
Proportion males in household	0.53 (0.17)	0.51 (0.18)	0.49 (0.18)	0.47 (0.15)
Head of household variables				
Age in years	37.73 (10.32)	39.51 (14.29)	38.86 (12.67)	38.02 (9.62)
Years of education	8.58 (2.77)	8.00 (3.59)	8.15 (3.95)	7.41 (3.92)
Gender (0=female, 1=male)	0.42 (0.50)	0.35 (0.48)	0.55 (0.50)	0.52 (0.50)
Farmer (0=no, 1=yes)	0.04 (0.20)	0.29 (0.46)	0.59 (0.50)	0.38 (0.49)
Salesperson (0=no, 1=yes)	0.50 (0.51)	0.45 (0.50)	0.01 (0.11)	0.21 (0.41)
Marital status (0=not married, 1=married)	0.69 (0.47)	0.68 (0.47)	0.86 (0.35)	0.84 (0.33)
Literacy (0=illiterate, 1=literacy)	1.00 (0.00)	0.96 (0.20)	0.70 (0.46)	0.79 (0.41)
Well-being variables				
Importance (1=not at all, 5=extremely)	4.88 (0.33)	4.88 (0.33)	4.04 (0.73)	4.02 (0.62)
Accomplishment (1=not at all, 5=extremely)	4.65 (0.75)	4.67 (0.72)	3.25 (1.15)	4.03 (0.57)
Children (1=not at all, 5=extremely)	4.65 (0.85)	4.57 (0.87)	3.99 (1.07)	4.21 (0.52)

Note: This table presents sample characteristics for households, heads of households and their well-being for urban areas and the three geographic areas. The values are the means/fractions of survey questions, with their standard deviations in brackets. Economic development is measured in terms of PPI^{TM} as is explained in Section 2.2.1. The urban/rural classification of Emwanu et al. (2007) is used to determine the proportion of households in urban areas. To measure well-being, three questions were asked to which the interviewees had to give an answer on a scale from 1 (”not at all”) to 5 (”extremely”). These questions were: ”How important do you believe you are to others?” (’Importance’); ”Do you believe you can accomplish in life what you want?” (’Accomplishment’); and ”Do you believe your children will succeed in education?” (’Children’). A ”mobile phone user” is defined as someone who *owns* a mobile phone.

Chapter 3

Keeping up with the Garcías: Social Comparisons and Subjective Well-Being in Bolivia

"Hay que redistribuir la riqueza, no la pobreza."
(You have to redistribute wealth, not poverty)

– Anonymous respondent

3.1 Introduction

The last years have seen a substantial increase in the number of studies dealing with subjective well-being (see MacKerron, 2011, for a comprehensive review).[1] Owing to this growing body of literature, we have learned much about the measurement of happiness, the determinants of happiness, and the use of happiness in formulating policy. In addition, the notion has now been established that happiness

[1] I will use the terms life satisfaction, happiness and subjective wellbeing interchangeably.

economics is "serious enough to be taken seriously" (Helliwell et al., 2012, p. 20). In the wake of the seminal Stiglitz-Sen-Fitoussi report of 2009, happiness has also attracted the attention of policymakers. Echoing the words of Stiglitz et al. (2009): "Measures of both objective and subjective wellbeing provide key information about peoples quality of life" [p.16]. If subjective wellbeing informs policy or even becomes a policy objective, it is essential to know what drives it.

A large strand of happiness research has focused on the determinants of subjective wellbeing (see Dolan et al., 2008, for a detailed review). A major result is the so-called 'Easterlin Paradox' regarding the relationship between income and happiness (Easterlin, 1974, 1995, 2005a,b). According to Easterlin, richer individuals are happier than poorer individuals *within* a country. However, richer countries do not necessarily have higher average happiness levels than their poorer counterparts. Furthermore, increasing income over time is not associated with a long-run increase in happiness. Recent literature has contested Easterlin's result, showing evidence of a link between subjective wellbeing and both GDP per capita and economic growth (Stevenson and Wolfers, 2008; Sacks et al., 2010).

Other important determinants of subjective wellbeing include but are not limited to age, education, marital status, and employment. Interestingly, happiness declines with age, but then picks up again between the age of forty and fifty (Blanchflower and Oswald, 2004, 2008; Hayo and Seifert, 2003; Helliwell, 2003; Helliwell et al., 2010).[2] Furthermore, the educated, married, and employed tend to report higher life satisfaction (Helliwell et al., 2012). The results with regards to gender are mixed: in developed countries women are happier than men, but this difference is smaller or even reversed in developing countries (Helliwell et al., 2012).

Despite the progress made in understanding the determinants of life satisfaction, the existing literature is limited in three respects. First, developing countries have been underrepresented. Even though cross-country studies increasingly include developing countries, more detailed studies at the country level are mostly confined to advanced economies. One example of an individual-level study is Knight

[2]This result has been coined the 'U-bend of life' by the Economist (The Economist, 2010b).

and Gunatilaka (2011), who analyse subjective wellbeing in China. As noted by MacKerron (2011), development economics is a surprising absentee when it comes to happiness research. The focus on developed countries can be explained by the lack of data from the developing world (Knight and Gunatilaka, 2011). However, results from rich countries cannot necessarily be generalised, as it is plausible that other factors influence happiness of poor countries' residents (Knight and Gunatilaka, 2011). For example, an increase in income might have a larger impact on a person's happiness if she is living at subsistence level than if all basic needs are met.

Second, few present studies include social reference groups, such as neighbours or colleagues. The phenomenon of using ones neighbours as a point of reference is commonly referred to as 'keeping up with the Joneses.' The 'relative deprivation hypothesis' implies that an individual feels worse off if her standards of living compare unfavourably with the people around her, the so-called 'reference group.' However, the reference income can also have a positive impact on subjective wellbeing if it signals what an individual might be able to attain in the future (Clark and D'Ambrosio, 2014).

It is difficult to determine how a reference group should be defined (Eibner and Evans, 2005) and data on a particular reference group is not always available. Therefore, reference groups are usually imposed, by defining them by geographic proximity, e.g. individuals from the same city, or similarity, in terms of e.g. age, gender, education (Clark and D'Ambrosio, 2014). Such studies have overwhelmingly found that the position within a reference group matters for subjective wellbeing (see Clark et al., 2008, for an extensive review and Clark and D'Ambrosio, 2014, for recent studies).

However, individuals probably compare themselves with 'social reference groups' - i.e. people whom they interact with frequently - rather than with such imposed reference groups. Using a large-scale survey, Clark and Senik (2010) show that Europeans are most likely to say that they compare themselves to colleagues. In contrast, Knight et al. (2009) find that fellow villagers are the most common comparison group in rural China. Nevertheless, studies on social reference groups are scarce. As Clark et al. (2008) points out: "despite the current abundance of microeconomic data, very few papers

have related individual wellbeing to co-workers wages. (...) The situation is equally sparse with respect to family and friends" [p. 108]. The few existing studies about social reference groups tend to use subjective comparisons, i.e. how people perceive their income relative to a reference group. In Clark and Senik (2010), individuals are asked to rank themselves among various reference groups, and they find that a more favourable ranking is associated with higher subjective wellbeing. Knight and Gunatilaka (2011) show that a higher perceived income - relative to fellow villagers - is positively correlated with happiness.

The third limitation of the present literature is the narrow focus on present life satisfaction, while an individual's expectation of her future happiness can also be informative. For example, someone who is optimistic about the future might be more productive or healthier. Therefore, it is also important to know the determinants of future life satisfaction. In particular, it is interesting to study the relation between the social reference group and future happiness, because it helps to further understand the role of reference groups in shaping aspirations.

This study adds to the literature on subjective wellbeing and reference groups by analysing the role of social reference groups in a developing country. Whereas most papers in the current development literature have focused on actual (objective) comparisons with a reference group (e.g. Fafchamps and Shilpi, 2008; Graham and Pettinato, 2002), I study perceived (subjective) comparisons. Another contribution is the analysis of future life satisfaction, in addition to present satisfaction. This paper is most closely related to Mangyo and Park (2011) who study the relation between social reference groups and self-reported health in China. They find that relatives and classmates are important reference groups for urban respondents, while neighbours are important in rural areas.

I use a unique cross-sectional data set from Bolivia. To my knowledge, this is the first study on determinants of subjective wellbeing in this country. Bolivia is known to be diverse, both in terms of demographics and geography, which offers desirable variation in my data. Fifty-five per cent of the Bolivian population is indigenous, income inequality is relatively high, and forty-five per cent of the

population lives on less than $2 a day.[3] Also, the recent development progress makes it an interesting case study. Since Evo Morales came to power in 2006, a new constitution has granted the indigenous population more rights and poverty has dropped by almost a half. Morales has reaped the benefits of the Bolivian natural gas boom and used revenues for social purposes, such as schools and hospitals (The Economist, 2014).

I study self-reported, and thus subjective, comparisons to several reference groups. An obvious complication in studying the relation between two subjective variables - in my case subjective well-being and subjective comparisons - is the potential omitted variable bias resulting from a generally positive or negative disposition of respondents. For example, if an individual is naturally optimistic, she will answer more favourably to both subjective questions than a respondent with a more pessimistic outlook. My econometric methodology, similar to Mangyo and Park (2011), deals with this omitted variable bias by estimating a variable that measures disposition. However, it does not deal with potential simultaneous causality of comparisons and life satisfaction. A two-way causal relationship is not unlikely, as happiness impacts factors that are important to relative standing, such as income (De Neve et al., 2013). I can therefore not draw any conclusions about causality.

I find a significant relation between subjective comparisons and life satisfaction. In particular, comparisons to family members and (former) classmates are important determinants of subjective wellbeing. I find that other reference groups are associated with future life satisfaction, namely colleagues and neighbours. More generally, I find that the drivers of life satisfaction are in line with present findings, except for marital status. In addition, future life satisfaction has different drivers than present life satisfaction. My econometric methodology allows for the correction of response bias through estimating a measure for optimism. I investigate and confirm the robustness of this method by using anchoring vignettes. Overall, I can conclude that both social reference groups and future life satisfaction provide additional valuable sources of information for subjective wellbeing research.

[3] All these statistics were taken from the CIA World Factbook: https://www.cia.gov/library/publications/the-world-factbook (last accessed on July 4, 2014). With a Gini index of 47, Bolivia has the 23^{rd} highest inequality worldwide.

The remainder of the paper is organised as follows. Section 3.2 discusses data collection and includes descriptive statistics. Section 3.3 outlines the methodology. Section 3.4 shows the results. Finally, Section 3.5 concludes.

3.2 Data

Data were collected by the author in January and February 2012. Oral surveys were done with 237 individuals in three departments[4] of Bolivia: Tarija (49.4% of the interviews), Santa Cruz (23.6%) and La Paz (27%). I chose these three departments, because they represent characteristically different parts of the country: *Los Valles*, *El Oriente* and *El Occidente*, respectively. The data set consists of three groups of variables: individual and household characteristics; comparisons; and subjective wellbeing measures. Descriptive statistics for these three categories are included in Panel A, B and C of Table 3.1, respectively.

[4]Bolivia is made up of nine departments, which are the main subdivisions of the country.

Table 3.1: Descriptive statistics

Variable	Mean	Std. Dev.	Min.	Max.	N
Panel A: Personal/household characteristics					
female	0.519	0.501	0	1	237
age (years)	39.627	13.308	18	78	236
education (years)	12.453	6.151	0	24	236
indigenous	0.251	0.435	0	1	235
able to read	0.903	0.297	0	1	237
able to write	0.895	0.308	0	1	237
head of household	0.820	0.385	0	1	233
married	0.608	0.489	0	1	237
lives in a rural area	0.22	0.415	0	1	236
per capita household income (Bolivianos/month)	817.781	555.773	71.776	2616.062	233
Panel B: Comparisons of standards of living					
comparison to family members (1=very low,...,5=very high)	3.122	0.938	1	5	237
comparison to classmates (1=very low,...,5=very high)	3.072	0.774	1	5	237
comparison to colleagues (1=very low,...,5=very high)	3.222	0.815	1	5	237
comparison to neighbours (1=very low,...,5=very high)	3.042	0.758	1	5	237
comparison to residents from city/town (1=very low,...,5=very high)	3.148	0.802	1	5	237
comparison to Bolivians (1=very low,...,5=very high)	3.101	1.037	1	5	237
autoposition (1=very low,...,10=very high)	4.212	1.974	1	9	236
Panel C: Subjective wellbeing (1=worst possible life,...,10=best possible life)					
present life satisfaction (1-10)	5.358	2.375	1	10	236
past life satisfaction (1-10)	4.506	2.528	1	10	237
future life satisfaction (1-10)	7.056	2.68	1	10	234

Note: Summary statistics of key variables. Unless specified otherwise, variables are binary that equal 1 if the respective condition is satisfied and 0 if not. Bolivianos is the Bolivian currency.

3.2.1 Individual and household characteristics

First, the data include individual characteristics (see Panel A in Table 3.1). The sample is gender balanced and the mean age of respondents is forty years. On average, the respondents have 12.5 years of education and around 90% can read and write. In my sample, 82% of the respondents are head of their household, 22% live in rural areas and 25% are indigenous.

Second, the data include variables that are also used on the Progress out of Poverty Index™ (PPI™) scorecard for Bolivia. It is difficult to measure economic development at the household level in developing countries, because clear income documentation is often not available. I therefore use the PPI, which was developed by the Grameen Foundation and can be used to determine the poverty likelihood of households. The technical details of Bolivia's index are outlined in Schreiner (2009). The index is calculated using the ten questions and scores outlined in Table 3.2. I use these variables as Schreiner (2009) shows that these are important determinants for poverty and thereby income.

Table 3.2 also includes the sample percentages for the answer categories of each PPI question. The household size is 4.5 on average. Of questions dealing with whether the household owns, has or uses a specific asset, the proportions range from 65% to 89%. About half of the households have at least one member in a blue- or white-collared job. Only four households (1.7%) had at least one child that did not attend the appropriate level and grade in school.

I use the PPI scorecard to get an assessment of the economic situation of the household and to estimate household income for my sample. The PPI for Bolivia is based on 4,148 households surveyed in the 2007 *Encuesta de Hogares* (EH) (Schreiner, 2009). I regress log per capita household income[5] on nine out of ten PPI questions using the EH data.[6] With the resulting regression coefficients I estimate per capita household income for my sample. The regression results are included in Table 3.A1. I find

[5] I calculate per capita household income by dividing household income from the EH data set by household size.

[6] I exclude question 10 concerning blue- and white-collared jobs, as 99% of observations are missing for this question in the EH data set. I use the frequency weights as given in the EH data set. See footnote 3 in Schreiner (2009), for more information on the weighting procedure.

an average per capita income of about 818 Bolivianos per month, which corresponds to 360 US dollars (PPP).[7] This income is slightly higher than the mean income of 772 Bolivianos in the EH sample. To my knowledge, this is the first time that household income is estimated using the PPI scorecard.

In general, my sample statistics differ from official statistics (see Table 3.3). For example, indigenous and rural residents are underrepresented in all departments. Also, women are underrepresented in Santa Cruz, whereas it is the opposite case in La Paz and Tarija. The sample mean education level is more than 1.5 times higher than the population's average in each department. Also, the sample is likely not to be representative for Bolivia as only three out of nine departments are included. Tarija is extremely overrepresented, with a proportion ten times higher than in the population. Therefore, one should be wary of generalising the results of this study to the whole of Bolivia. However, the internal validity of my conclusions are not in jeopardy.

3.2.2 Comparisons

Next, the data include variables that reflect subjective comparisons of the respondent with specific reference groups (see Panel B in Table 3.1). Respondents were asked to compare their standards of living to those of six different groups: (extended) family members, (former) classmates with the same level of schooling as the respondent, colleagues, neighbours, people from the same city/village, and Bolivians.[8] For each group, they were asked whether their standards were much worse (=1), slightly

[7] I use the exchange rate of 0.142 Boliviano/US Dollar, as recorded on January 30, 2012, which is halfway my data collection period. The CPI for Bolivia was 113.364 in 2007 (the year of the EH) and 157.236 in 2012 (see http://data.worldbank.org/data-catalog/world-development-indicators, last accessed on June 25, 2014). The calculation is as follows: $817.781 * (1 + (157.236 - 113.364)/113.364) * 0.142 = 161.065$ 2012 US Dollars per capita per month. I need to multiply this amount by 1/0.448 to get the PPP equivalent, since the PPP conversion factor (GDP) to market exchange rate was 0.448 in 2012 ((see http://data.worldbank.org/data-catalog/world-development-indicators, last accessed on June 25, 2014).

[8] This question is a slight alteration of the following question used in the China Inequality and Distributive Justice survey project, conducted in Fall 2004: "Compared with the average living standard of [your relatives, classmates with the same level of schooling as you, your coworkers, your neighbours, others in the same county or city, others in the same province, others living in China], do you feel your living standard is much better, a little better, about the same, a little worse, or much worse?" (Mangyo and Park, 2011). In this study, I excluded the option "in the same province."

Table 3.2: Progress out of Poverty™ scorecard for Bolivia with sample proportions

Question	Answer	Points	Sample %
1. How many household members are there?	A: Seven or more	0	13.9
	B: Six	7	13.5
	C: Five	11	16.5
	D: Four	16	19.4
	E: Three	17	15.2
	F: Two	26	15.2
	G: One	35	6.3
2. How many household members ages 6 to 17 currently attend school at the level and grade that they enrolled in for this calendar year?	A: Not all	0	1.7
	B: All	2	47.3
	C: No children ages 6 to 17	4	51.1
3 What is the main construction material of the floors of the residence?	A: Earth, bricks, or other	0	21.5
	B: Wooden planks, cement, hardwood floors, parquet, rugs or carpets	4	47.7
	C: Tile (mosaic, stone, or ceramic)	10	30.8
4. What is the main fuel used for cooking?	A: Firewood, dung/manure, kerosene, LPG in a cylinder, or other	0	64.5
	B: Piped-in natural gas, electricity, or does not cook	7	35.5
5. Does the household own, have, or use a refrigerator or freezer?	A: No	0	28.3
	B: Yes	5	71.7
6. Does the household own, have or use a dining-room set (table and chairs)?	A: No	0	28.4
	B: Yes	5	71.6
7. Does the household own, have or use a television?	A: No	0	11.4
	B: Yes	10	88.6
8. Does the household own, have or use a VCR or DVD player?	A: No	0	19.4
	B: Yes	6	80.6
9. Does the household own, have or use a stereo or hi-fi system?	A: No	0	35.4
	B: Yes	5	64.6
10. Are any household members employed in blue-collar or white-collar jobs?	A: No	0	47.5
	B: Yes	13	52.5

Note: The Progress out of Poverty Index™ (PPI ™) for Bolivia is based on the ten questions included in this table with the points for each answer given in the third column. The PPI™ is the sum of points corresponding to the respondent's answers. This scorecard is introduced in Schreiner (2009). The fourth column indicates the percentage of the sample that gave the respective answer.

Table 3.3: Sample statistics [official statistics] per department

	Tarija	**Santa Cruz**	**La Paz**
# interviews	119	56	64
% of total sample	49.6 [4.7]	23.5 [24.5]	26.9 [28.4]
% rural	26.5 [36.7]	16.1 [23.8]	18.8 [33.9]
% women	51.7 [50.1]	46.4 [49.5]	56.3 [50.4]
% indigenous	11.2 [22]	23.2 [40]	51.6 [82]
mean age of adults (years)	37.6 [39.3]	43.1 [37.4]	40.1 [41.1]
mean education (years)	12.0 [7.0]	13.4 [8.1]	12.6 [7.9]

Note: This table shows summary statistics per department. For comparison, official statistics are given in square brackets. Official statistics on '% indigenous' come from *Los pueblos indigenas de Bolivia* by CEPAL (2005). Mean age of adults (18 years and older) is based on own calculations using 2007 *Encuesta de Hogares*. Other official statistics come from the 2001 census and are provided by the *Instituto Nacional de Estadistica* (INE); data were downloaded from www.ine.gob.bo, last accessed on 12 April 2013.

worse (=2), the same (=3), slightly better (=4) or much better (=5). I imputed missing values with value 3 and include a dummy for individuals with missing values in all regressions in this paper. The average score for all these measures lies somewhat above 3, meaning that the average respondent thinks she is slightly better off than the individuals in the respective reference group. The correlations among the comparison measures are positive and significant, and lie between 0.14 and 0.50.

Also, the survey includes a question regarding autoposition, in which the respondent is asked to position herself on the national income distribution. I ask the following: "In our society, there are groups that tend to be towards the top and groups that tend to be towards the bottom. Below is a scale that runs from top to bottom. Where would you put yourself and your family[9] on this scale?"[10] In the surveys, I use a figure depicting a line of people and a scale to facilitate the question (see Figure 3.1). The scale runs from 1 to 10. The average respondent positions herself at 4.2, which is more than 1

[9]Here, 'family' refers to nuclear family.

[10]This question was used in the International Social Survey Program (ISSP) 1999 survey. See http://zacat.gesis.org/webview/index.jsp?object=http://zacat.gesis.org/obj/fStudy/ZA3430. Last accessed on 20 December 2011.

point below the midpoint of 5.5. No respondent puts herself in the top position of 10.

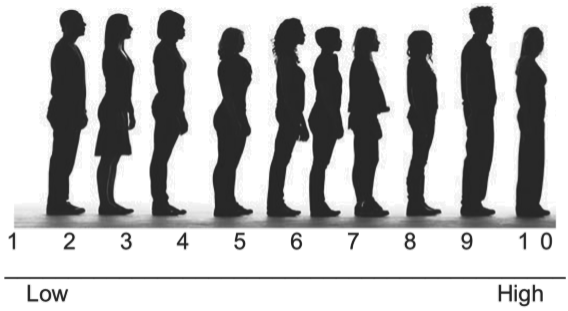


Figure 3.1: Image used in questionnaire to facilitate autoposition question (translated from the Spanish original)

3.2.3 Subjective wellbeing

Last, the data include subjective wellbeing variables (see Panel C in Table 3.1). I use the "self-anchoring striving scale" as introduced by Cantril (1965), often referred to as 'Cantril's ladder.' This measure is, among others, used by the Gallup World Poll and it currently covers the widest range of countries (Helliwell et al., 2012). The question reads as follows:

> Suppose the top of the ladder represents the best possible life for you, and the bottom the worst possible life for you, where on the ladder do you feel you personally stand at this time? (Cantril, 1965)

Similar to Cantril (1965) I use a scale from 1 (worst possible life) to 10 (best possible life). In the surveys, I use a picture of a ladder to facilitate the question (see Figure 3.2). I will refer to this variable as 'present life satisfaction.' In line with the Gallup World Poll, I also asked where the respondent believes she will be five years from now ('future life satisfaction') and where she was five years ago ('past life

satisfaction'). On average, there is an increasing trend going from past to present to future, with means 4.5, 5.4 and 7.1, respectively.[11] The histogram in Figure 3.3[12] shows a bell-curved shape.[13] Figure 3.4, however, shows that future life satisfaction is far from normally distributed.[14] The strong negative skew can be explained by a large proportion of respondents expecting extremely high satisfaction in the future. Sixty-five respondents, 27.8% of the sample, have a score of ten on future life satisfaction.

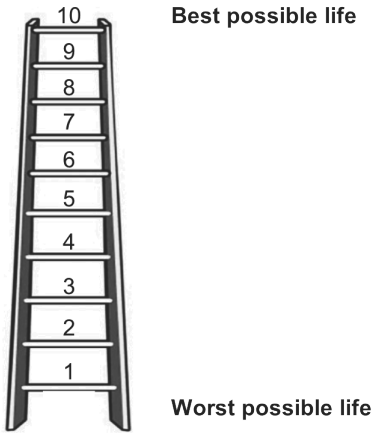


Figure 3.2: Image used in questionnaire to facilitate life satisfaction question (translated from the Spanish original)

[11]Gallup combines present and future life satisfaction to assign respondents to the categories 'thriving' (present satisfaction>6 and future satisfaction>7), 'suffering' (present<5 and future<5) and 'struggling' (any other combination of scores). In my sample, 31% are thriving, 51% are struggling, and 18% are suffering. In the Gallup data for Bolivia, these percentages are 28%, 65%, and 7%, respectively. See http://www.gallup.com/poll/world.aspx?ref=b and http://www.gallup.com/poll/122453/Understanding-Gallup-Uses-Cantril-Scale.aspx (last accessed on August 6, 2014).

[12]A few respondents indicated that their satisfaction was between two values (e.g. "Between 4 and 5"), which explains 'dips' in the histogram in between the integers.

[13]The *sktest* in Stata does not reject the null of a skewness of 0 (p=0.723), but does reject the null of a kurtosis of 3 (p=0.002). Altogether, normality is rejected (p=0.013). The *sktest* is similar to the Jarque-Bera test (Jarque and Bera, 1987), but adjusts the statistic for sample size as proposed by Royston (1991) and D'Agostino et al. (1990).

[14]The *sktest* in Stata rejects the null of a skewness of 0 (p=0.001) and the null of a kurtosis of 3 (p=0.000). Altogether, normality is rejected (p=0.000). See the previous footnote for more information on the test statistics used.

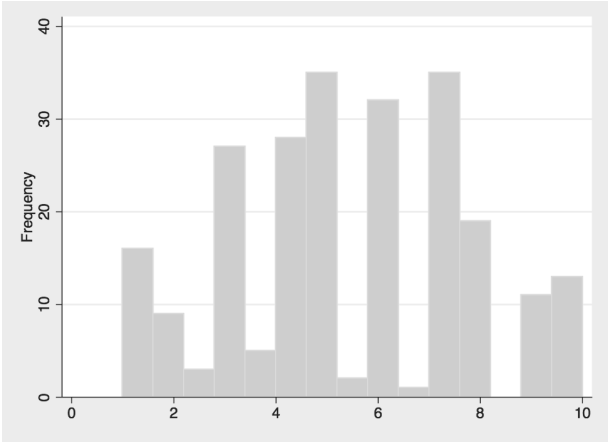


Figure 3.3: Present life satisfaction

3.3 Methodology

If I ignore non-material sources of wellbeing, subjective wellbeing can be seen as a function of standards of living and material aspirations. Aspirations determine what an individual defines as a '1' (worst life possible) and a '10' (best life possible) on the life satisfaction scale. She bases the anchoring of this scale on a reference group, such as her colleagues or neighbours, as well as on her past levels of material wellbeing. Her happiness level follows from comparing her own living standards to her subjective scale, and is thus related to what she (subconsciously) defines as her best and worst possible life. An increase in her living standards does not necessarily imply higher life satisfaction, as aspirations might have changed as a result of the changing levels of material comfort of herself or her surroundings. It follows from this rationale that happiness is affected by absolute as well as relative living standards. In this study, I approximate standards of living with per capita household income, which I expect to have a positive relation with happiness.

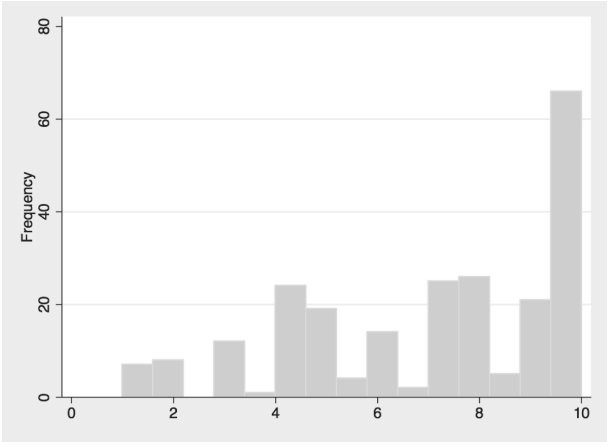


Figure 3.4: Future life satisfaction

I aim to investigate how happiness depends on the comparison with the reference group. Ideally, I wish to estimate the following regression

$$satisfaction_i = \alpha + X_i\beta + \gamma comparison_i + \delta logincome_i + \epsilon_i, \tag{3.1}$$

where $satisfaction_i$ is a subjective wellbeing variable for individual i, matrix X_i includes individual characteristics and $logincome_i$ is log per capita income of i's household. $comparison_i$ reflects how respondent i compares her living standards to those of a specific reference group, e.g. her neighbours. The higher the value of this variable, the more favourable the respondent compares her living standard to those of the reference group; the variable ranges from 'much worse' to 'much better.' I expect that income increases life satisfaction ($\delta > 0$). All else being equal, I believe a more favourable comparison to increase life satisfaction ($\gamma > 0$). As was mentioned before, the subjective happiness scale of a respondent is, among others, determined by the distribution of living standards of her reference group.

For example, a successful uncle might become the example of a 'good life' and a poor cousin might represent a 'bad life.' Therefore, the higher the comparison with the members of this group, the closer the respondent is to her best possible life - in the words of Cantril (1965) - and the higher is the life satisfaction.

However, there is a potential endogeneity problem when estimating Equation 3.1 as both *satisfaction* and *comparison* are influenced by response bias. For example, a more optimistic individual will give a higher satisfaction *and* a more favourable subjective comparison, *ceteris paribus*, than someone with a more negative outlook. Hence, the model in Equation 3.1 suffers from an omitted variable bias as a consequence of leaving out 'optimism', or 'outlook' as I will call it in the remainder of this paper.

In order to deal with the endogeneity in my model, I estimate the response bias and include it in the original model. I ask the respondent to position herself and her family on the national income distribution ($autoposition$, see Section 3.2.2). I assume that the answer to this question is determined entirely by three factors: the household income, a geographic bias, and the (unobserved) outlook bias. The geographic bias matters because it is difficult for an individual to oversee the income distribution of her entire country. Hence, a respondent relies on what she sees in her closer surroundings. More formally,

$$autoposition_i = \alpha + \sum_{d=1}^{D}\sum_{k=1}^{K} \beta_{dk} * D_{idk} + \gamma * logincome_i + \eta_i, \tag{3.2}$$

where $autoposition_i$ reflects where individual i locates herself on the national income distribution and D_{idk} is a geographic dummy that is equal to 1 if individual i lives in department d and city size k.[15] I fix $\beta_{11} = 0$ for identification. Given my assumptions, the response bias of individual i should be equal

[15]Department $d \in (1, 2, 3)$, with departments Tarija, Santa Cruz and La Paz respectively. City size $k \in (1, 2, 3, 4, 5)$ with size classification 'city', 'suburb', 'town', 'village' and 'farm house/countryside', respectively.

to the error term η_i in equation 3.2. So,

$$outlook_i = autoposition_i - a - \sum_{d=1}^{D}\sum_{k=1}^{K} b_{dk} * D_{idk} - c * logincome_i, \tag{3.3}$$

where a, b_{dk}, and c are the coefficient estimates of α, β_{dk}, and γ in equation 3.2 respectively. The variable $outlook_i$ is the outlook of individual i, where a higher value reflects a more positive outlook. All else being equal, individuals with a more positive outlook are expected to report a higher level of life satisfaction. This procedure resembles the Heckman error correction, as it also uses a generated regressor to deal with endogeneity bias (Heckman, 1978).

After estimating the outlook bias with ordinary least squares (OLS) estimation of Equation 3.2, I include it in Equation 3.1. Now the outlook bias is added, there is no omitted variable bias. Formally, I thus estimate

$$satisfaction_i = \alpha + X_i\beta + \gamma comparison_i + \delta logincome_i + \zeta \widehat{outlook_i} + \epsilon_i. \tag{3.4}$$

I calculate Huber-White heteroscedasticity robust standard errors. Since $outlook_i$ is a generated regressor, OLS estimation does not lead to the correct standard error for this coefficient (Pagan, 1984). I therefore calculate the correct error with a bootstrap procedure. More specifically, I use a nonparametric bootstrap (Wooldridge, 2002) with 1,000 replications.[16]

This methodology is based on Mangyo and Park (2011), but is different in two main respects. First, Mangyo and Park (2011) use the respective comparison variable in Equation 3.2 instead of $autoposition$. Hence, their measure of the outlook variable depends on the reference group used. I prefer to use a different subjective question to estimate outlook, because I wish to obtain an outlook measure that is independent of the answer to the comparison question and is the same across comparisons. Second, Mangyo and Park (2011) instrument the income variable because it is likely to suffer

[16] I executed the bootstrap estimation in Stata 11.2.

from measurement error. As I do not have an appropriate instrument, I choose not to use instrumental variables.[17]

3.4 Results

3.4.1 Main results

I calculate the outlook variable using OLS estimation of Equation 3.3 (see Table 3.A2). Figure 3.5 shows the correlation between outlook and selected variables. Outlook is positively and significantly correlated with education, while women and indigenous respondents have a significantly more negative outlook. By construction (see Equation 3.3), outlook is uncorrelated with income. In addition, I do not find a significant correlation with age or marital status.

The OLS estimation[18] of Equation 3.4 for present life satisfaction is shown in Table 3.4.[19] In the first six regressions, I include one comparison variable at a time. The coefficient of log income is positive and significant: a 1% increase in income rises life satisfaction with 1.0 to 1.4 points. Three comparisons are significant and positive: family members, classmates, and Bolivians. A 1-category increase in comparison (e.g. moving from 'slightly worse' to 'the same') with family members, classmates or Bolivians has a similar effect as a 0.2%, 0.4% or 0.2% increase in income, respectively. None of the comparisons remains significant if I finally include all comparison variables in regression 7.

Moreover, I find that women, the higher educated, and respondents from Santa Cruz report a significantly higher life satisfaction. To illustrate, the effect of being a woman is the same as a 0.5% increase

[17] If measurement error and income are positively correlated, the income coefficient is likely to be overestimated. As a consequence, the variance in outlook would be lower in case of an unbiased estimate and ordering of individuals based on outlook might change.

[18] I do not estimate ordered probit as the answer categories are too many and too small, because I also allowed for half points in life satisfaction. For example, someone who answered "between 5 and 6", received a score of 5.5.

[19] See Table 3.A3 for the results without the outlook variable. In general, the same coefficients are significant, but are lower. One exception is female, which is insignificant when excluding the outlook variable. When excluding the outlook, the coefficients on comparisons are higher and more often significantly different from zero.

in income, and the effect of living in Santa Cruz is about the same as a 1% increase in income.[20] The education effect is quite small economically, as one would need 12 years of additional education to experience an 1-point increase in happiness. The coefficient of age squared is significantly positive, indicating a u-shaped relationship between age and life satisfaction. However, the coefficient is small: one would need to get older or younger by 22.4 years[21] in order to increase happiness by one point. As expected, outlook is significant and positive. I do not find a significant effect for being married, living in La Paz, living in a rural area, and being indigenous.

The results for future life satisfaction are included in Table 3.5.[22] Future satisfaction can, for a large part, be explained by present satisfaction. If present satisfaction increases by 1 point, future satisfaction increases by about 0.7 points, *ceteris paribus*. This is not a surprising result, seeing that the correlation between these two variables is 0.77. Log income is significantly positive in all regressions: a 1% increase in income is associated with a 0.5-point increase in future happiness. Comparisons with colleagues and neighbours matter: if a respondent rates herself higher among these groups, her future life satisfaction is likely to be higher. To illustrate, the effect of a 1-category increase in comparison with colleagues is comparable to a 0.6% increase in income. This effect is 0.7% for a 1-category increase in comparison with neighbours. Only the comparison with neighbours remains significant when I include all comparisons in one regression.

Age is significantly and negatively correlated with future life satisfaction: an increase in age of 24 years is associated with a 1-point drop in happiness. Respondents living in La Paz report that their future happiness is 1.1 points higher, *ceteris paribus*, than those in other departments. Outlook is not significant anymore; its effect is likely captured by present life satisfaction.

In the Data section I showed that the distribution of future satisfaction was negatively skewed. This was a consequence of the high number of respondents that answered they expect to attain the maximum

[20] I use the estimates from model 7 for calculations in this paragraph.

[21] $x = 22.4$ follows from solving $0.002x^2 = 1$.

[22] See Table 3.A4 for the results without the outlook variable. Results are similar, because the outlook variable did not have a significant coefficient.

happiness in the future. If I exclude these respondents and re-estimate the model for future satisfaction, the estimated coefficients closely resemble those in the original model (see Table 3.6).

3.4.2 Robustness

I collected additional data to investigate the robustness of my outlook measure. In 2013, 115 individuals were interviewed in the department of Tarija, using a similar questionnaire to the 2012 survey. I added so-called "anchoring vignettes" (King et al., 2004) to the questionnaire in order to obtain an alternative measure of outlook. Vignettes are short descriptions of hypothetical individuals. After reading the vignette to the respondent, he or she is asked to assess the life satisfaction of the imaginary person. She rates the happiness of the vignette person on the same scale that she used to answer the question about her own life satisfaction, i.e. a ten-point scale. Usually, vignettes are used to make corrections for heterogeneity in response response styles, in my case with regards to happiness. This method relies on two key assumptions (King et al., 2004): *response consistency* - individuals use the ten-point scale of the happiness question in the same way for themselves as for the hypothetical persons; and *vignette equivalence* - vignettes represent the same level of happiness to all respondents.

Appendix 3.B includes the vignettes; the descriptive statistics of the 2013 sample are shown in Table 3.A5. It is noteworthy that vignettes for male and female respondents are generally the same, only the gender of the hypothetical figure is matched to the respondent's gender. This is common practice in the vignette literature (see e.g. Hopkins and King, 2010). In line with Ravallion (2012), the vignettes are ordered in increasing level of (expected) happiness.[23] Following Kapteyn et al. (2009), each vignette relates to four domains that are known to be important to happiness (Easterlin and Sawangfa, 2007): family/friends, job, income, and health.

Someone who relatively overestimates (underestimates) vignette happiness is likely to also overes-

[23] I estimate the outlook variable for 2013 in a similar way as I did for 2012. The only difference is that I have less geographic variation. I only have data for one department, and only know whether the respondent lives in a rural or urban area. Instead of the geographic dummy, I now include a dummy for whether the respondent lives in an urban area.

timate (underestimate) her own happiness due to a relatively optimistic (pessimistic) outlook. Hence, if my outlook measure indeed measures whether a person is optimistic or pessimistic, it should have a positive correlation with the vignette scores. Figure 3.6 confirms this hypothesis. The correlation of outlook with vignette score equals 0.118 (p-value 0.267), 0.173 (p-value 0.104), 0.210 (p-value 0.046), and 0.010 (0.925), respectively. The low level of variation in scores for vignette 4 can be explained by the fact that all the four domains were highly positive. There was less room for bias as this person was 'objectively' happy. In short, I show that vignette scores and outlook are positively correlated, which supports the validity of my outlook measure. However, it should be noted that the correlations are not strong and only once significant (at the 5% level).

3.5 Conclusion

I study the relationship between happiness and subjective comparisons to social reference groups. Using unique cross-sectional data from Bolivia, I find that social reference groups are important for both present and future life satisfaction. If individuals rate themselves higher among these groups, their life satisfaction is likely to be higher. My econometric model handles the likely omitted variable bias by estimating respondents' optimistic disposition. However, the model does not deal with potential simultaneous causality resulting from an impact of happiness on comparisons. I can therefore not draw any conclusions about causality. Results from anchoring vignettes support the validity of my estimation of the respondents' disposition. However, this evidence is weak and needs more research.

More specifically, family members and classmates matter for present satisfaction. These groups tend to be fairly stable, i.e. members are not likely to leave the group. In contrast, future life satisfaction is correlated with comparisons to colleagues and neighbours, which are subject to change over the course of an individual's life. These findings suggest that stable reference groups may be more informative for a 'stock variable' such as present satisfaction, whereas dynamic reference groups influ-

ence a 'flow variable' such as the expected change in life satisfaction over the coming five years. This explanation is suggestive rather than conclusive and should be investigated further.

The general finding that social reference groups are important for subjective wellbeing is in line with present literature. More specifically, the importance of relatives and classmates concurs with Mangyo and Park (2011), who find that these groups matter for subjective health in urban areas in China. That I find similar reference groups as in the Chinese urban population, rather than the rural population, can be explained by the fact that 78% of my sample lives in cities. To my knowledge, there are no studies relating future happiness to social reference groups, so I cannot compare my conclusions on future happiness to previous work.

More generally, this study looked at the determinants of subjective wellbeing. To my knowledge this is the first study of this kind in Bolivia. I find that subjective wellbeing increases with income and education, which is in line with current research. Also, the u-shaped relationship between age and happiness has been found in previous studies. Women in my sample tend to be happier than men, which is in accordance with evidence for developed countries. In the developing world, however, results with regards to gender have been mixed. Even though there is widespread evidence of a positive association between marriage and happiness, this study did not find a significant relation.

This study adds to the evidence base on future life satisfaction, which has rarely been studied in happiness literature. For a large part, future life satisfaction can be explained by present life satisfaction. In addition, log income is positively associated with future happiness, controlling for present life satisfaction. This is not surprising, since the better endowed are probably more able to realise their aspirations in the future. In addition, future satisfaction decreases with age, which can probably partly be account for by an expected deterioration in health. Also, I find that residents of the department of La Paz were more optimistic. This finding might be a consequence of the recent rise of President Evo Morales who started an 'indigenous revolution', granting more rights to indigenous Bolivians. Since La Paz has a relatively high proportion of indigenous individuals, this area might have experienced a

relative increase in optimism as a result of Morales' policies. Furthermore, indigenous respondents in other areas indicated in conversations that Morales has mainly focused on the indigenous population in the Andean area, which includes La Paz. Though merely anecdotal, this offers evidence why I did not find a significant impact for indigenous respondents in general.

The happiness literature would benefit from an increased focus on developing countries, trying to understand cross-country differences in determinants. In particular, the channels behind the relationship between social reference groups and subjective wellbeing merit further research. It would be interesting to understand how a comparison comes about, for example whether it is based merely on material possessions or on other aspects of the reference groups. Also, both the dynamic dimension of the relationship and causality deserve attention. Future data collection endeavours would merit including questions on future satisfaction and anchoring vignettes in order to correct for response bias. This paper is, to my knowledge, the first to use the PPI scorecard to assess household income. Future studies in developing countries would benefit from using poverty scorecards, as they enable researchers to quickly evaluate income where income documentation is not readily available.

The policy implications of my results on social comparisons are not straightforward. One might use my findings to argue that inequality should be reduced. On the other hand, if relative deprivation is a fundamentally human trait, then some individuals will always feel better or worse off than the people around them. In that case, subjective wellbeing analyses should control for these comparisons, but policy should not try to influence them. Subjective comparisons might simply be a fact of life.

Table 3.4: OLS estimation for present life satisfaction

	1	2	3	4	5	6	7
constant	-3.736**	-4.309***	-3.633**	-3.874***	-3.739***	-3.587**	-4.911***
	(1.45)	(1.49)	(1.46)	(1.42)	(1.43)	(1.43)	(1.50)
log income	1.112***	1.086***	1.125***	1.135***	1.207***	1.081***	1.054***
	(0.24)	(0.24)	(0.24)	(0.25)	(0.25)	(0.25)	(0.25)
lives in Santa Cruz	1.055***	1.144***	1.208***	1.154***	1.189***	1.065***	1.068***
	(0.34)	(0.34)	(0.34)	(0.34)	(0.34)	(0.34)	(0.35)
lives in La Paz	0.117	0.204	0.183	0.154	0.186	0.089	0.305
	(0.29)	(0.29)	(0.29)	(0.29)	(0.30)	(0.29)	(0.29)
married	-0.215	-0.202	-0.222	-0.203	-0.153	-0.173	-0.023
	(0.26)	(0.26)	(0.26)	(0.26)	(0.25)	(0.26)	(0.27)
female	0.491**	0.582**	0.539**	0.467*	0.462*	0.543**	0.552**
	(0.25)	(0.24)	(0.26)	(0.24)	(0.25)	(0.25)	(0.26)
indigenous	0.041	0.004	-0.029	-0.006	0.003	0.026	0.079
	(0.33)	(0.33)	(0.32)	(0.33)	(0.34)	(0.33)	(0.33)
rural	0.066	0.194	0.138	0.075	0.173	0.209	0.147
	(0.40)	(0.39)	(0.40)	(0.40)	(0.40)	(0.41)	(0.41)
age (years)	-0.014	-0.014	-0.015	-0.016	-0.018	-0.018	-0.020
	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
age squared	0.002**	0.002**	0.002**	0.002**	0.002**	0.002**	0.002**
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
education (years)	0.082**	0.086**	0.078**	0.077**	0.087**	0.079**	0.083**
	(0.04)	(0.04)	(0.04)	(0.04)	(0.04)	(0.04)	(0.04)
outlook	0.530***	0.495***	0.514***	0.504***	0.516***	0.471***	0.448***
	(0.08)	(0.09)	(0.08)	(0.09)	(0.09)	(0.09)	(0.10)
comparisons (1=much worse off,...,5=much better off)							
family members	0.210*						0.140
	(0.13)						(0.13)
classmates		0.392**					0.278
		(0.18)					(0.19)
colleagues			0.156				0.135
			(0.16)				(0.17)
neighbours				0.246			0.168
				(0.16)			(0.16)
residents from city/town					0.003		-0.208
					(0.18)		(0.19)
Bolivians						0.267**	0.199
						(0.13)	(0.14)
Observations	226	226	226	226	226	226	226
R-squared	0.487	0.494	0.484	0.485	0.487	0.489	0.531

Note: OLS estimation for present life satisfaction (1-10). Heteroscedasticity robust standard errors are given in parentheses. Bootstrapped errors are given for the variable 'outlook.' *** indicates significance at the 0.01 level; ** indicates significance at the 0.05 level; * indicates significance at the 0.10 level. Dummies for missing variables for comparisons are included in the original regression, but the coefficients are not shown in this table.

Table 3.5: OLS estimation for future life satisfaction, controlled for present life satisfaction

	1	2	3	4	5	6	7
constant	0.436	0.231	0.127	-0.105	0.467	0.329	-0.437
	(1.44)	(1.47)	(1.47)	(1.50)	(1.45)	(1.44)	(1.47)
log income	0.559**	0.558**	0.481**	0.525**	0.530**	0.561**	0.493**
	(0.23)	(0.23)	(0.22)	(0.22)	(0.22)	(0.23)	(0.22)
present life satisfaction (1-10)	0.732***	0.726***	0.725***	0.715***	0.728***	0.728***	0.742***
	(0.06)	(0.06)	(0.06)	(0.06)	(0.06)	(0.06)	(0.06)
lives in Santa Cruz	0.571	0.533	0.522	0.556	0.540	0.475	0.511
	(0.36)	(0.36)	(0.35)	(0.35)	(0.36)	(0.36)	(0.37)
lives in La Paz	1.060***	1.065***	1.055***	1.057***	0.996***	1.036***	1.128***
	(0.32)	(0.32)	(0.33)	(0.31)	(0.32)	(0.31)	(0.33)
married	-0.013	0.047	-0.016	0.028	0.008	0.002	0.011
	(0.25)	(0.26)	(0.25)	(0.25)	(0.25)	(0.25)	(0.24)
female	0.143	0.141	0.093	0.145	0.181	0.218	0.100
	(0.23)	(0.24)	(0.24)	(0.23)	(0.24)	(0.23)	(0.26)
indigenous	0.026	0.086	0.039	0.049	0.074	0.062	-0.015
	(0.32)	(0.32)	(0.33)	(0.31)	(0.32)	(0.31)	(0.32)
rural	-0.098	-0.060	-0.199	-0.227	-0.157	-0.158	-0.162
	(0.34)	(0.33)	(0.33)	(0.32)	(0.34)	(0.33)	(0.31)
age (years)	-0.042***	-0.046***	-0.042***	-0.043***	-0.043***	-0.043***	-0.047***
	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
age squared	0.000	0.000	-0.000	0.000	0.000	-0.000	-0.000
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
education (years)	0.008	0.026	0.015	0.003	0.008	0.013	0.020
	(0.03)	(0.03)	(0.03)	(0.03)	(0.03)	(0.03)	(0.03)
outlook	0.065	0.065	0.039	0.051	0.066	0.068	0.003
	(0.08)	(0.08)	(0.08)	(0.08)	(0.09)	(0.08)	(0.09)
comparisons (1=much worse off,...,5=much better off)							
family members	0.031						-0.036
	(0.11)						(0.12)
classmates		0.053					-0.036
		(0.13)					(0.15)
colleagues			0.283**				0.195
			(0.14)				(0.14)
neighbours				0.357***			0.305*
				(0.14)			(0.16)
residents from city/town					0.094		-0.016
					(0.16)		(0.19)
Bolivians						0.059	0.033
						(0.13)	(0.16)
Observations	223	223	223	223	223	223	223
R-squared	0.667	0.669	0.673	0.674	0.665	0.672	0.694

Note: OLS estimation for future life satisfaction (1-10). Heteroscedasticity robust standard errors are given in parentheses. Bootstrapped errors are given for the variable 'outlook.' *** indicates significance at the 0.01 level; ** indicates significance at the 0.05 level; * indicates significance at the 0.10 level. Dummies for missing variables for comparisons are included in the original regression, but the coefficients are not shown in this table.

Table 3.6: OLS estimation for future life satisfaction, controlled for present life satisfaction (excluding maximum happiness)

	1	2	3	4	5	6	7
constant	0.453	0.832	0.105	0.013	0.674	0.390	0.187
	(1.61)	(1.56)	(1.66)	(1.66)	(1.61)	(1.62)	(1.58)
log income	0.565**	0.576**	0.432	0.503*	0.641**	0.572**	0.510*
	(0.27)	(0.27)	(0.27)	(0.27)	(0.28)	(0.28)	(0.28)
present life satisfaction (1-10)	0.730***	0.729***	0.733***	0.734***	0.700***	0.725***	0.755***
	(0.07)	(0.07)	(0.07)	(0.06)	(0.08)	(0.07)	(0.07)
lives in Santa Cruz	0.421	0.351	0.328	0.254	0.381	0.249	0.343
	(0.45)	(0.45)	(0.45)	(0.47)	(0.44)	(0.46)	(0.47)
lives in La Paz	0.944***	0.887***	0.813**	0.855***	0.986***	0.883***	1.037***
	(0.31)	(0.31)	(0.32)	(0.30)	(0.32)	(0.31)	(0.32)
married	0.053	0.131	0.089	0.118	0.123	0.065	0.076
	(0.26)	(0.26)	(0.25)	(0.26)	(0.25)	(0.26)	(0.25)
female	-0.436*	-0.507*	-0.388	-0.361	-0.445*	-0.319	-0.514*
	(0.24)	(0.26)	(0.25)	(0.24)	(0.26)	(0.24)	(0.30)
indigenous	0.009	0.074	0.116	0.105	-0.000	0.084	-0.047
	(0.33)	(0.33)	(0.33)	(0.32)	(0.33)	(0.32)	(0.33)
rural	0.064	0.042	-0.025	-0.029	0.075	-0.054	0.110
	(0.33)	(0.33)	(0.34)	(0.31)	(0.33)	(0.32)	(0.31)
age (years)	-0.042***	-0.044***	-0.038***	-0.043***	-0.041***	-0.042***	-0.043***
	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
age squared	0.000	-0.000	-0.000	0.000	-0.000	-0.000	-0.000
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
education (years)	-0.011	0.002	0.003	-0.010	-0.004	-0.002	-0.005
	(0.03)	(0.03)	(0.03)	(0.03)	(0.03)	(0.03)	(0.03)
outlook	-0.005	0.028	-0.007	-0.024	0.058	0.025	-0.062
	(0.10)	(0.11)	(0.11)	(0.09)	(0.11)	(0.11)	(0.12)
comparisons (1=much worse off,...,5=much better off)							
family members	0.029						-0.009
	(0.12)						(0.12)
classmates		-0.150					-0.258
		(0.16)					(0.19)
colleagues			0.318*				0.263
			(0.17)				(0.18)
neighbours				0.309**			0.326*
				(0.14)			(0.18)
residents from city/town					-0.195		-0.231
					(0.21)		(0.22)
Bolivians						0.019	0.095
						(0.14)	(0.16)
Observations	160	160	160	160	160	160	160
R-squared	0.657	0.653	0.656	0.659	0.651	0.661	0.694

Note: OLS estimation for future life satisfaction (1-10), excluding respondents who answered that their future happiness would be 10. Heteroscedasticity robust standard errors are given in parentheses. Bootstrapped errors are given for the variable 'outlook.' *** indicates significance at the 0.01 level; ** indicates significance at the 0.05 level; * indicates significance at the 0.10 level. Dummies for missing variables for comparisons are included in the original regression, but the coefficients are not shown in this table.

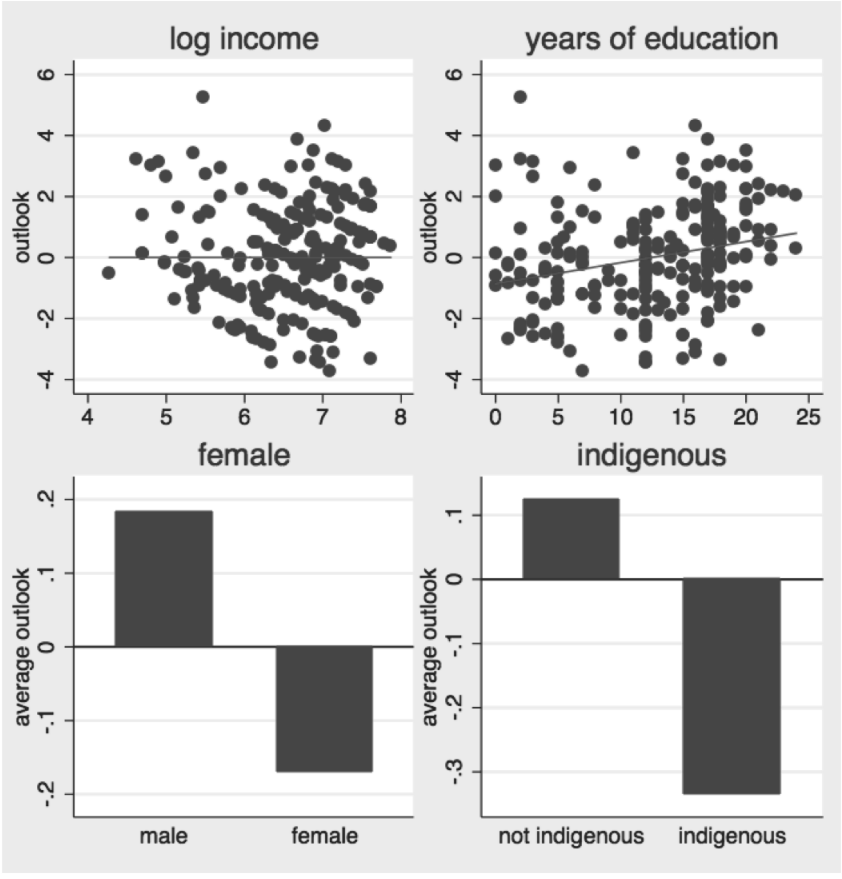


Figure 3.5: Scatter plots of the outlook variable and selected variables

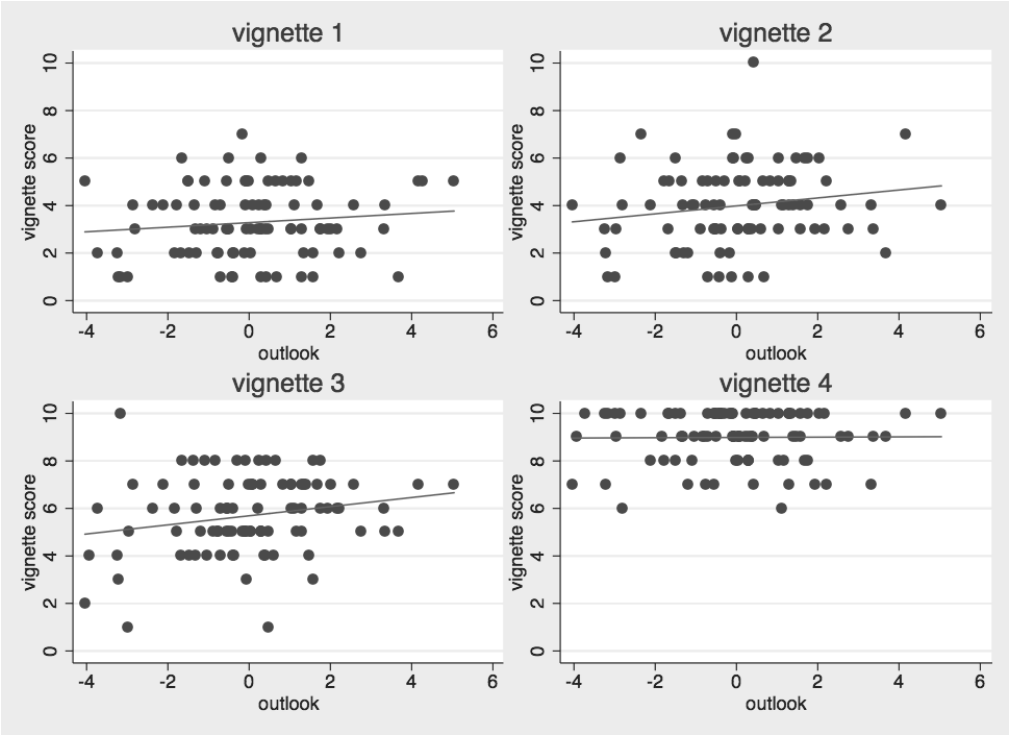


Figure 3.6: Scatter plots of vignette scores and outlook (2013 sample)

3.A Appendix tables

Table 3.A1: OLS estimation for log per capita household income (EH data)

constant	5.659***
	(0.01)
household size	-0.086***
	(0.00)
education: all children ages 6-17 in household go to school	0.210***
	(0.01)
education: no children ages 6-17 in household	0.580***
	(0.01)
floor material: wooden planks, cement, hardwood floors, parquet, rugs or carpets	-0.317***
	(0.00)
floor material: earth, bricks, or other	-0.574***
	(0.00)
cooking fuel: firewood, dung/manure, kerosene, LPG in a cylinder, or other	-0.329***
	(0.00)
household owns, has or uses a refrigerator or freezer	0.378***
	(0.00)
household owns, has or uses a dining-room set	0.281***
	(0.00)
household owns, has or uses a television	0.703***
	(0.00)
household owns, has or uses a VCR or DVD player	0.193***
	(0.00)
household owns, has or uses a stereo or hi-fi	0.161***
	(0.00)
Obs	1553188
R-squared	0.426

Note: Data from the 2007 *Encuesta de Hogares* (EH) are used to estimate the coefficients in this table. EH frequency weights are used. The variables are selected using the Progress out of Poverty™ scorecard from Schreiner (2009). All variables, except for household size, are binary and have value 1 if the respective condition is satisfied. For the categorical variable 'education' the omitted category is "not all children ages 6-17 in school"; for 'floor material' the omitted category is "tile (mosaic, stone, or ceramic)"; for 'cooking fuel' the omitted category is "piped-in natural gas, electricity, or does not cook." Standard errors are given in parentheses. *** indicates significance at the 0.01 level; ** indicates significance at the 0.05 level; * indicates significance at the 0.10 level.

Table 3.A2: OLS estimation for autoposition

	1
constant	-3.346***
	(1.25)
log income	1.140***
	(0.18)
department/city size dummies	
14	-0.155
	(0.40)
15	1.241
	(1.74)
21	0.542
	(0.43)
22	-1.837
	(1.74)
23	0.825**
	(0.40)
24	-1.633
	(1.76)
25	-0.091
	(0.68)
31	-0.153
	(0.32)
34	-0.394
	(0.63)
35	0.090
	(1.25)
41	1.966
	(1.74)
51	-0.834
	(1.74)
Observations	231
R-squared	0.261

Note: OLS estimation of Equation 3.2. Standard errors are given in parentheses. The first number of department/city size dummy reflects the department number $d \in (1, 2, 3)$, with departments Tarija, Santa Cruz and La Paz respectively. The second number is city size $k \in (1, 2, 3, 4, 5)$ with size classification 'city', 'suburb', 'town', 'village' and 'farm house/countryside', respectively. The dummy 11 is omitted and thus serves as the baseline. *** indicates significance at the 0.01 level; ** indicates significance at the 0.05 level; * indicates significance at the 0.10 level.

Table 3.A3: OLS estimation for present life satisfaction (excluding outlook variable)

	1	2	3	4	5	6	7
constant	-2.816	-3.775**	-2.964*	-3.326**	-3.106*	-2.859*	-4.974***
	(1.74)	(1.68)	(1.71)	(1.63)	(1.68)	(1.65)	(1.66)
log income	0.840***	0.792***	0.786***	0.855***	0.868***	0.763***	0.754***
	(0.27)	(0.27)	(0.27)	(0.27)	(0.28)	(0.27)	(0.26)
lives in Santa Cruz	0.892**	0.952**	1.022***	0.964**	0.999***	0.806**	0.908**
	(0.38)	(0.37)	(0.38)	(0.37)	(0.37)	(0.36)	(0.37)
lives in La Paz	0.124	0.209	0.174	0.147	0.069	0.001	0.298
	(0.34)	(0.32)	(0.32)	(0.33)	(0.34)	(0.32)	(0.32)
married	-0.091	-0.038	-0.102	-0.054	0.010	-0.002	0.136
	(0.29)	(0.28)	(0.29)	(0.28)	(0.27)	(0.28)	(0.28)
female	0.364	0.519*	0.400	0.363	0.428	0.531*	0.496*
	(0.27)	(0.27)	(0.28)	(0.27)	(0.27)	(0.27)	(0.28)
indigenous	-0.114	-0.087	-0.157	-0.124	-0.037	-0.034	0.032
	(0.37)	(0.36)	(0.35)	(0.37)	(0.37)	(0.36)	(0.35)
rural	0.229	0.374	0.207	0.158	0.174	0.391	0.246
	(0.45)	(0.43)	(0.44)	(0.45)	(0.44)	(0.44)	(0.43)
age (years)	-0.017	-0.018	-0.017	-0.019	-0.025*	-0.024*	-0.029**
	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
age squared	0.002***	0.002***	0.002**	0.002***	0.002***	0.002**	0.002**
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
education (years)	0.143***	0.155***	0.141***	0.130***	0.137***	0.124***	0.120***
	(0.04)	(0.04)	(0.04)	(0.04)	(0.04)	(0.04)	(0.04)
comparisons (1=much worse off,...,5=much better off)							
family members	0.273*						0.063
	(0.14)						(0.14)
classmates		0.580***					0.278
		(0.20)					(0.22)
colleagues			0.427**				0.299
			(0.18)				(0.18)
neighbours				0.483***			0.288*
				(0.16)			(0.15)
residents from city/town					0.370*		-0.053
					(0.21)		(0.22)
Bolivians						0.569***	0.406***
						(0.12)	(0.14)
Observations	227	227	227	227	227	227	227
R-squared	0.369	0.392	0.378	0.382	0.384	0.406	0.467

Note: OLS estimation for present life satisfaction (1-10), excluding the outlook variable. Estimations are shown without including the outlook variable. Heteroscedasticity robust standard errors are given in parentheses. *** indicates significance at the 0.01 level; ** indicates significance at the 0.05 level; * indicates significance at the 0.10 level. Dummies for missing variables for comparisons are included in the original regression, but the coefficients are not shown in this table.

Table 3.A4: OLS estimation for future life satisfaction, controlled for present life satisfaction (excluding outlook variable)

	1	2	3	4	5	6	7
constant	0.553	0.341	0.155	-0.062	0.545	0.440	-0.437
	(1.41)	(1.44)	(1.44)	(1.47)	(1.42)	(1.39)	(1.47)
log income	0.511**	0.505**	0.451**	0.489**	0.476**	0.505**	0.493**
	(0.21)	(0.22)	(0.20)	(0.21)	(0.20)	(0.21)	(0.22)
present life satisfaction (1-10)	0.756***	0.749***	0.739***	0.733***	0.750***	0.749***	0.742***
	(0.06)	(0.06)	(0.05)	(0.05)	(0.06)	(0.06)	(0.06)
lives in Santa Cruz	0.509	0.472	0.474	0.499	0.473	0.404	0.511
	(0.34)	(0.34)	(0.34)	(0.34)	(0.34)	(0.34)	(0.37)
lives in La Paz	1.058***	1.062***	1.051***	1.053***	0.980***	1.023***	1.128***
	(0.32)	(0.32)	(0.33)	(0.31)	(0.32)	(0.31)	(0.33)
married	0.020	0.083	0.009	0.060	0.046	0.040	0.011
	(0.25)	(0.25)	(0.24)	(0.24)	(0.24)	(0.24)	(0.24)
female	0.114	0.116	0.071	0.124	0.162	0.201	0.100
	(0.23)	(0.23)	(0.23)	(0.22)	(0.24)	(0.23)	(0.26)
indigenous	0.017	0.081	0.041	0.048	0.078	0.061	-0.015
	(0.32)	(0.32)	(0.33)	(0.31)	(0.32)	(0.31)	(0.32)
rural	-0.075	-0.037	-0.189	-0.211	-0.151	-0.132	-0.162
	(0.34)	(0.33)	(0.33)	(0.31)	(0.34)	(0.33)	(0.31)
age (years)	-0.042***	-0.046***	-0.042***	-0.043***	-0.043***	-0.044***	-0.047***
	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
age squared	0.000	0.000	-0.000	0.000	0.000	-0.000	-0.000
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
education (years)	0.013	0.033	0.019	0.007	0.012	0.018	0.020
	(0.03)	(0.03)	(0.03)	(0.03)	(0.03)	(0.03)	(0.03)
comparisons (1=much worse off,...,5=much better off)							
family members	0.034						-0.036
	(0.11)						(0.12)
classmates		0.064					-0.036
		(0.12)					(0.15)
colleagues			0.297**				0.195
			(0.13)				(0.14)
neighbours				0.371***			0.305*
				(0.13)			(0.16)
residents from city/town					0.131		-0.016
					(0.15)		(0.19)
Bolivians						0.088	0.033
						(0.12)	(0.16)
outlook							0.003
							(0.09)
Observations	224	224	224	224	224	224	223
R-squared	0.666	0.668	0.673	0.674	0.665	0.671	0.694

Note: OLS estimation for future life satisfaction (1-10), excluding the outlook variable. Estimations are shown without including the outlook variable. Heteroscedasticity robust standard errors are given in parentheses. *** indicates significance at the 0.01 level; ** indicates significance at the 0.05 level; * indicates significance at the 0.10 level. Dummies for missing variables for comparisons are included in the original regression, but the coefficients are not shown in this table.

Table 3.A5: Descriptive statistics for 2013 sample

Variable	Mean	Std. Dev.	Min.	Max.	N
Panel A: Personal/household characteristics					
female	0.46	0.501	0	1	113
age (years)	39.684	14.02	18	81	114
education (years)	12.315	6.033	0	22	111
indigenous	0.098	0.299	0	1	112
able to read	0.964	0.186	0	1	112
head of household	0.678	0.469	0	1	115
lives in a rural area	0.217	0.414	0	1	115
per capita household income (Bolivianos/month)	1065.129	677.475	64.308	2773.343	112
Panel B: Comparisons of standards of living					
comparison to family members (1=very low,...,5=very high)	3.113	0.758	1	5	115
comparison to classmates (1=very low,...,5=very high)	3.209	0.778	1	5	115
comparison to colleagues (1=very low,...,5=very high)	3.217	0.659	2	5	115
comparison to neighbours (1=very low,...,5=very high)	3.148	0.728	1	5	115
comparison to residents from city/town (1=very low,...,5=very high)	3.043	0.921	1	5	115
comparison to Bolivians (1=very low,...,5=very high)	3.130	0.923	1	5	115
autoposition (1=very low,...,10=very high)	5.990	1.914	1	10	99
Panel C: Subjective wellbeing (1=worst possible life,...,10=best possible life)					
present life satisfaction (1-10)	7.342	1.791	2	10	111
past life satisfaction (1-10)	6.864	2.052	2	10	110
future life satisfaction (1-10)	8.255	1.876	1	10	102
vignette 1	3.352	1.513	1	8	105
vignette 2	4.029	1.640	1	10	104
vignette 3	5.769	1.604	1	10	108
vignette 4	9.045	1.080	6	10	108

Note: Summary statistics of key variables. Unless specified otherwise, variables are binary that equal 1 if the respective condition is satisfied and 0 if not. Bolivianos is the Bolivian currency. The exchange rate was 0.142 Boliviano/US Dollar on January 30, 2012.

3.B Anchoring vignettes

These vignettes are translated from the Spanish versions in the original questionnaire.

3.B.1 Vignettes for female respondents

1. Sonia

Sonia is separated and never gets to see her children. She does not have any close friends. She makes her money with agriculture. In a good month, she earns 500 Bolivianos. She has serious heart problems and gets tired easily. On a scale from 1 to 10, how satisfied do you believe Sonia is with her life in general?

2. Martha

Martha is divorced and has to take care of one child. She sells vegetables at the market. She makes 1000 Bolivianos a month. She often cannot sleep because of pain in her knees, elbows, wrists and fingers. On a scale from 1 to 10, how satisfied do you believe Martha is with her life in general?

3. Marina

Marina is married with two children. She fights sometimes with her husband because he is jealous. She has a group of good friends, which she has known since her childhood. The family income is 3000 Bolivianos a month. Marina works as a teacher. She is healthy, but she has stomach pains every now and then. On a scale from 1 to 10, how satisfied do you believe Marina is with her life in general?

4. Angela

Angela is happily married and has three children. All her children do well in school. The family income is 9000 Bolivianos per month. She likes her job and has a fixed contract. In her spare time, she likes to do sports. On a scale from 1 to 10, how satisfied do you believe Angela is with her life in general?

3.B.2 Vignettes for male respondents

1. Jorge

Jorge is separated and never gets to see his children. He does not have any close friends. He makes his money with agriculture. In a good month, he earns 500 Bolivianos. He has serious heart problems and gets tired easily. On a scale from 1 to 10, how satisfied do you believe Jorge is with his life in general?

2. Luis

Luis is divorced and has to take care of one child. He sells vegetables at the market. He makes 1000 Bolivianos a month. He often cannot sleep because of pain in his knees, elbows, wrists and fingers. On a scale from 1 to 10, how satisfied do you believe Luis is with his life in general?

3. Mario

Mario is married with two children. He fights sometimes with his wife because she is jealous. He has a group of good friends, which he has known since her childhood. The family income is 3000 Bolivianos a month. Mario works as a teacher. He is healthy, but he has stomach pains every now and then. On a scale from 1 to 10, how satisfied do you believe Mario is with his life in general?

4. Carlos

Carlos is happily married and has three children. All his children do well in school. The family income is 9000 Bolivianos per month. He likes his job and has a fixed contract. In his spare time, he likes to do sports. On a scale from 1 to 10, how satisfied do you believe Carlos is with his life in general?

Chapter 4

Birds of a Feather: The Spread of Happiness and Behavior in a Dutch Sorority

4.1 Introduction

The maxim "birds of a feather flock together" dates back to the 16^{th} century. It confirms what we all observe in our daily lives: social bonds tend to exist between individuals with similar characteristics, background or behaviour. Do similar people simply seek each other out (*homophily*) or do they also become more similar because of their interaction (*induction*)? The disentanglement of these two effects is important for policymakers. For example, if unhealthy behaviour is subject to induction, a government can use this 'contagion' effect in the design of an effective campaign. This paper aims to disentangle homophily and induction by using the unique sorting process of a Dutch sorority, while controlling for confounding factors. In particular, I look at happiness,[1] behaviour, and educational outcomes.

[1] I use the terms life satisfaction, happiness and subjective wellbeing interchangeably.

4.1.1 Happiness, fraternities and peer effects

This study lies at the intersection of three strands of research. The first strand is literature on happiness and social connections. There is strong evidence of the importance of social relationships for happiness (see Dolan et al., 2008, for a detailed overview), but little is known about how happiness spreads among individuals. A notable exception is Fowler and Christakis (2008) who study the diffusion of happiness in a large social network. They find that happy and unhappy people cluster together and that this association can partly be explained by induction. Interestingly, they find that correlation in happiness extends up to three degrees of separation (for example, the friend of a friend's friend).

The second strand is research regarding fraternal organisations. Fraternal organisations are social clubs of university members. Male-only and female-only organisations are called 'fraternities' and 'sororities', respectively. Studies in this field focus mainly on the effects of membership and are confined to American colleges. Naturally, self-selection plays a role in such studies. After all, it might be that individuals who choose to become member of a fraternal organisation are different from those who choose not to.

Many studies focus only on correlations between membership and outcomes, but some try to disentangle these selection effects from the effects of membership. DeSimone (2007) and DeSimone (2009) find a positive impact of membership on alcohol (ab)use. Eating disorders and body mass index (BMI) are also affected by membership (Averett et al., 2013). More generally, Gilbert and Meyer (2004) find that self-selected groups of young women have more similar eating behaviour and depression levels. In addition, Scott-Sheldon et al. (2008) show a higher propensity to smoke and have more sexual partners among members of a fraternal organisation, but causality is not investigated. Studies using social network analysis can shed more light on the development of behaviour in a fraternal organisation, but such studies are rare. Exceptions are Phua (2011) who show that smoking and drinking diffuses through the college fraternity network and Reifman et al. (2006) who study college drinking more generally.

Beyond unhealthy behaviour, fraternity/sorority membership is correlated to cheating behaviour

(Kerkvliet, 1994) and lower academic performance (Grubb, 2006; DeBard et al., 2006). Pascarella et al. (2001) confirm that this lower performance is a consequence of membership. In contrast to these findings, Pike (2000) find that membership is related to gains in cognitive development. Membership is also associated with greater social capital (Green and Brock, 2005) and leadership (Martin et al., 2012), but the causal direction is unclear. Furthermore, students who use fraternity/sorority network connections in their job search are more likely to get a high-paid job (Marmaros and Sacerdote, 2002).

The third strand of literature uses experimental settings to analyse peer effects. In such studies, individuals are randomly assigned to peers, which allows to control for selection effects. Chetty et al. (2011) study Project STAR, a programme that randomly assigns children to classrooms, and show that children with better classmates in kindergarten have better long-term outcomes, such as higher earnings. On the other hand, Angrist and Lang (2004) show that the Metco programme, a project that aims to desegregate schools, had a modest and impermanent impact on outcomes. Kling et al. (2007) exploit another US policy experiment, Moving to Opportunity, to show that moving to a better neighbourhood has great mental health benefits for adults and female youth but no effect on adult self-sufficiency or physical health. Guryan et al. (2009) exploit the setup of professional golf tournaments to analyse peer effects at the workplace and find no evidence for effects on performance.

In terms of content, Sacerdote (2001) is most closely related to my paper. In Dartmouth college, dorm roommates were randomly assigned to each other. Sacerdote (2001) uses this random allocation to study peer effects among college roommates. He finds significant effects on grade point average and on the decision to join social groups like a fraternity. For other outcomes, such as the choice of college major, peer effects are absent. The paper concludes that "[a] fruitful area of future research would be to examine similar data in other educational settings" [p.703]. In terms of methodology, my paper resembles Fowler and Christakis (2008) who study the spread of happiness in a large social network. This methodology has been criticised by several authors (e.g. Cohen-Cole and Fletcher, 2008) and I will further address their concerns in the methodology section of this paper.

4.1.2 Fraternal organisations in The Netherlands

Student life in The Netherlands started with the opening of the first university in Leiden in 1575 (Huf-man et al., 2007). The first fraternal organisations were established at the beginning of the 19th century. Nowadays, there are multiple fraternal organisations in Dutch university towns, differing in history, traditions and religious background (Dronkers and Hillege, 1998). These organisations are generally larger than their American counterparts.

A specific example of a fraternity is the 'corps' (plural 'corpora'), the oldest and most traditional fraternal organisation in every university town. Since the university system in The Netherlands is relatively egalitarian, "[t]he only institutions which enabled elite distinction at the universities were the student fraternities (studenten-corpora), which have played an important role in the creation of the Dutch old boys network" (Heemskerk and Fennema, 2009, p. 813). Dronkers and Hillege (1998) find that corps membership is important for attaining different types of elite positions. To my knowledge, Dronkers and Hillege (1998) is the only quantitative study regarding fraternal organisations in the Netherlands.

There is some data availability on fraternal organisations in The Netherlands. As opposed to studies looking at American fraternal organisations, Gruijter et al. (2005) find that students in student associations in Leiden have a higher chance of getting their first year degree and a lower chance of dropping out. Maalsté (2000) writes that members of student associations in The Netherlands drink more than non-affiliated students (23 versus 13 glasses a week).

4.1.3 The spread of happiness and behaviour in a Dutch sorority

In this paper, I study peer effects of happiness, behaviour, and study outcomes in a Dutch sorority. The unique setup of this sorority enables me to disentangle homophily from induction. I collect data at the moment new members enter the sorority, before they get to know each other. After a few weeks, they

are expected to form so-called 'year clubs' (henceforth 'clubs'), which are social groups within each cohort. I control for homophily using baseline data and for confounding factors using the relationships between fellow club members.

My contributions are twofold. First, my study is one of the few network studies related to happiness and, to my knowledge, the first to study happiness in the setting of a fraternal organisation. As far as I know, Fowler and Christakis (2008) is the only other network study on happiness. In their study, social connections are formed at different points in time and it is not always clear when a social tie was created. In my paper, the social network is formed at a fixed point in time, which makes it possible to collect data before the formation starts. Second, it is one of the few network studies on fraternal organisations and one of the first quantitative studies on Dutch fraternal organisations. Current literature on fraternal organisations mainly focuses on the outcomes of membership, but studies do not always account for selection effects nor look at the dynamics of outcome variables.

Next to happiness, I look at other variables of interest, such as BMI, alcohol use, and study performance. I choose these variables because they relate to recurrent themes in the literature on fraternal organisations. I do not find evidence for homophily in happiness, behaviour or educational outcomes. In other words, individuals who are similar in these respects are not more likely to flock together. However, parental income and existing networks are important in the sorting process: respondents tend to be connected if they both have rich families, come from the same city, both live in a sorority house or follow the same study major. Moreover, I find that happiness, alcohol use, and grades are subject to induction - i.e. contagion - whereas correlation in BMI and relationship status is due to confounding factors. Different checks confirm the robustness of these effects. Since the induction effects are mainly driven by mutual friendships, they are possibly a consequence of homophily of friendships.

The remainder of this paper is structured as follows. Section 4.2 describes the sorority and the data. Section 4.3 outlines the methodology for the analysis of homophily and induction. Section 4.4 presents the results, including robustness checks, and interprets the estimated effects. Section 4.5 discusses the

results, comparing them to previous studies, and concludes.

4.2 Data

4.2.1 Setting

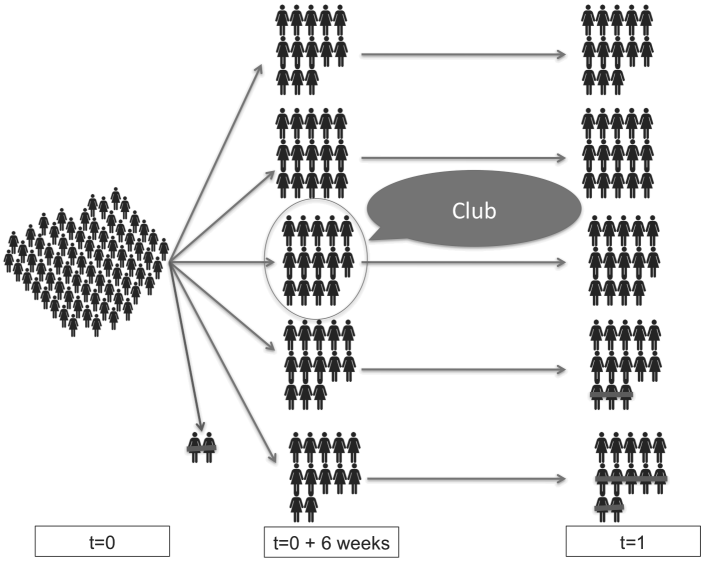
Data were collected at a Dutch sorority, which is one of the few all-female fraternal organisations in The Netherlands. The sorority undertakes a wide range of activities, such as social gatherings, charity events, and sports activities. Sorority members cluster in 'clubs', which consist of members of the same cohort. Clubs are important components of the sorority's social network: fellow club members (henceforth 'clubmates') dine together on a weekly basis, jointly carry out mandatory assignments for the sorority in the first year, and vacation together. Individuals tend to be close to their clubmates, even after sorority membership officially ends.[2] An individual cannot be member of more than one club. This paper focuses on clubs, since (i) they are the most important social groups in the sorority, (ii) all members enter the club at the same moment, and (iii) members usually do not know each other before the club is formed.

See Figure 4.1 for a schematic overview of the club sorting process. New members enter the sorority at the beginning of the academic year. After the first six weeks of their membership they are expected to form clubs. The sorting process is endogenous, but has some exogenous elements: the board of the sorority decides upon the maximum size of the club and the maximum number of clubs within the cohort. I collected data at the beginning of the academic year (t=0), immediately after individuals officially become sorority members. By doing so, I obtain baseline measures of happiness and other variables of interest. It should be noted that individuals have to pass an introduction period in order to become a member of the sorority, during which there is some opportunity for social interaction. However, these opportunities are so scarce that I assume they do not affect my results. A second round

[2] Usually, members leave the sorority after five years of membership.

of data collection is done after one year (t=1), after clubs have formed and clubmates have interacted for almost a full year. As depicted in Figure 4.1, some members drop out during the year, either before or after the club formation.

Figure 4.1: Graphic of club sorting process



Data were collected in Fall 2012 and Fall 2013 using questionnaires. In 2012, I conducted a questionnaire among the new members (2012 cohort). In 2013, a follow-up survey was done among the 2012 cohort. Also, the new cohort (2013 cohort) was interviewed. As an additional data source, I use the 'Almanac,' the yearbook of the sorority that lists all members.

4.2.2 Baseline data

The 2012 cohort (N=69) is divided in five clubs, with on average 13 members. Members of the 2013 cohort (N=101) are split up in seven clubs, with 14 members on average. Table 4.1 shows the baseline descriptive statistics for the 2012 and 2013 cohorts together.[3] The response rates were 99% for both cohorts.[4] The sample is clearly homogenous in terms of background: most of the respondents come from Dutch, high-income, and highly educated families. Almost 90% attend the highest level of secondary education and the most popular study major is business. In addition, 43% live in a sorority house, i.e. housing exclusively meant for sorority members. This proportion varies considerably across clubs: it is more than 90% for three clubs, while three clubs have less than 10% of its members living in a sorority house.

Happiness is measured on an eleven-point scale using 'Cantril's ladder.'[5] This measure is, among others, used by the Gallup World Poll, whose happiness data currently cover the widest range of countries (Helliwell et al., 2012). I will refer to this variable as 'happiness.' In line with Gallup, I ask how happy the respondent believes she will be in the future ('future happiness'), specifically one, five, and ten years from now. Also, I ask how happy she was five years ago ('past happiness'). The average happiness is 7.4 and respondents foresee an increasing trend in their happiness from past to future.[6]

[3] Appendix Tables 4.A1 and 4.A2 show separate descriptive statistics for the two cohorts separately. Differences between the samples are generally not statistically significant, which I tested with a t-test. Individuals in the 2013 cohort tend to study business or economics less often, and less respondents attend the highest level of secondary education. Also, happiness and future happiness (1 and 10 years) is lower in the 2013 cohort. Appendix Tables 4.A3 and 4.A4 show summary statistics for the separate clubs of the 2012 and 2013 cohorts, respectively. All within-cohort differences between clubs are statistically insignificant (Bonferroni multiple-comparison test), except for one: the age difference between two clubs in the 2012 cohort is significant at the 10% level.

[4] This exceptionally high response rate can be explained by the fact that members can easily be traced back with the help of the board of the sorority and the Almanac. Members usually do not move away from their college town until after five years.

[5] The "self-anchoring striving scale" (Cantril, 1965) is often referred to as 'Cantril's ladder.' The question reads as follows: "Suppose the top of the ladder represents the best possible life for you, and the bottom the worst possible life for you, where on the ladder do you feel you personally stand at this time?" (Cantril, 1965).

[6] Gallup combines present and future life satisfaction to assign respondents to the categories 'thriving' (present satisfaction>6 and future satisfaction>7), 'suffering' (present<5 and future<5) and 'struggling' (any other combination

In addition, I ask about the importance that respondents assign to certain aspects of life.[7] All respondents find family and friends important; more than three-quarters find sports and partying important; almost forty per cent find politics important, and only six per cent find religion important. Nearly half of the respondents define their health as very good or excellent and seventy per cent do sports at least once a week. Body mass index (BMI) is an indicator for weight relative to height;[8] a 'healthy weight' is indicated by a BMI score between twenty and twenty-five. The average sample BMI is twenty-one. Most respondents drink alcohol at least twice a week and three-quarters report to drink more than five glasses during a night on which alcohol is consumed ('drinking night').[9] Almost one in five reported never to have had sexual intercourse, while about a quarter have had four or more bed partners in their life.

4.2.3 Panel data for 2012 cohort

Table 4.2 shows the summary statistics of the outcome variables for the 2012 cohort as measured in 2012 and 2013. Note that, different from Table 4.1, I now report the 'raw' measures of the categorical variables (see the table note for more information on the scales of variables). Respondents drink significantly more in 2013 than in 2012 and their number of bed partners increased significantly. There is no significant change on average in happiness, sports, BMI or relationship status. Although the proportion of respondents in a serious relationship is identical, six respondents (11%) switched from being single

of scores). In my sample, 84% are thriving, 16% are struggling, and none are suffering. In the Dutch population, these percentages are 61%, 38%, and 1%, respectively. Hence, my sample is on average happier in terms of the Gallup classification. See http://www.gallup.com/poll/world.aspx?ref=b and http://www.gallup.com/poll/122453/Understanding-Gallup-Uses-Cantril-Scale.aspx (last accessed on August 4, 2014).

[7] These questions were taken from World Values Survey Wave 6. I transformed the answers to a binary variable indicating whether the respondent found an aspect important (answer categories "very important" or "rather important") or not ("not very important" or "not at all important"). See http://www.worldvaluessurvey.org/WVSOnline.jsp (last accessed on July 2, 2014).

[8] The formula for BMI is $\frac{weight(kg)}{(height(m))^2}$.

[9] These numbers amount to between 15 and 19 glasses per week on average. Nieuwenhuis and Postmes (2011) describe that fraternity members in the Dutch university town Groningen drink 21 glasses on average. Since this number includes men, who drink more than women, the Rotterdam respondents seem to be in line with their Groningen counterparts.

Table 4.1: Summary statistics baseline variables 2012 and 2013 cohort

Variable	Mean	Std. Dev.	N
age	18.553	0.877	170
born in The Netherlands	0.918	0.276	170
mother born in The Netherlands	0.888	0.316	170
father born in The Netherlands	0.905	0.294	169
parents married	0.771	0.422	170
higher than average household income	0.917	0.276	169
mother highly educated	0.749	0.435	167
father highly educated	0.842	0.365	165
highest high school level	0.806	0.397	170
grade final exam (1-10)	7.119	0.733	168
studies economics	0.071	0.258	168
studies business	0.333	0.473	168
studies medicine	0.095	0.294	168
follows the highest level of secondary education	0.876	0.331	169
lives in a sorority house	0.424	0.496	170
happiness	7.408	0.954	169
future happiness (1yr)	8.141	0.756	170
future happiness (5yr)	8.359	0.758	170
future happiness (10yr)	8.518	0.937	170
past happiness (5yr)	7.107	1.414	169
family important	1	0	169
friends important	1	0	169
religion important	0.06	0.237	168
sports important	0.787	0.411	169
party important	0.893	0.31	168
politics important	0.375	0.486	168
very good or excellent health	0.458	0.5	168
does sports at least once a week	0.697	0.461	165
weight (kg)	61.487	6.59	167
height (cm)	170.799	6.265	169
BMI	21.113	2.108	167
alcohol at least twice a week	0.905	0.294	168
alcohol at least 5 glasses a night	0.732	0.444	168
heterosexual	0.994	0.077	169
now in serious relationship	0.274	0.447	168
had at least one serious relationship	0.653	0.478	167
no bed partners	0.184	0.388	158
4 or more bed partners	0.234	0.425	158

in 2012 to being in a relationship in 2013 and seven respondents (12%) vice versa. This switching is more or less equally distributed among clubs.[10] In the follow-up survey, I also collected data on the number of study credits (ECTS) and the average grade obtained in the previous year. Students receive a specific number of ECTS after passing a university course and they are required to obtain sixty ECTS in order to complete one year of university. The average respondent did not obtain the total amount of sixty ECTS and scored an average grade of 6.7 on a ten-point scale.

The difference in sample size is due entirely to attrition of respondents who quit the sorority. Conditional on current membership, the response rate for the follow-up survey was 100%. Tables 4.3 and 4.4 show baseline summary statistics for the 2012 cohort, split up according to whether an individual is currently member or not. The average respondent who left the sorority has, compared to current members, a higher final exam grade, studies business more often, has lower present and future happiness, finds religion more important, and drinks alcohol less frequently. Since I want to compare the effect of peers' behaviour, the attrition does not affect my conclusions. After all, these respondents cease to be 'peers' after quitting the sorority according to my definition. It should be noted, however, that the attrition took place in only two clubs: one club lost three and another lost seven members. I include the attrited individuals in the analysis of homophily, but not in the estimations for induction.

[10] In club 1, 3 respondents switch from being in a relationship to being single (2 respondents vice versa). These numbers for clubs 2, 3, 4, and 5 are 4(2), 0(1), 0(1) and 0(0), respectively.

Table 4.2: Summary statistics for 2012 cohort in 2012 and 2013

	yr2012	yr2013
happiness (0=very unhappy,...,10=very happy)	7.62	7.74
	(0.90)	(0.97)
happiness future (1yr)	8.29	8.26
	(0.75)	(0.74)
happiness future (5yr)	8.45	8.31
	(0.74)	(0.86)
happiness future (10yr)	8.68	8.79
	(0.81)	(0.91)
happiness past (5yr)	7.00	6.64
	(1.38)	(1.37)
sports frequency (1=never,...,6=more than 5 times per week)	3.68	3.84
	(1.01)	(1.04)
BMI	21.11	21.54
	(2.20)	(2.39)
alcohol quantity (1=never,...,6=10 or more glasses)	4.00	4.62***
	(1.03)	(1.06)
now in serious relationship	0.28	0.28
	(0.45)	(0.45)
bed partners (1=0,...,5=10 or more)	1.16	1.98***
	(1.05)	(1.28)
ECTS previous year (1=0-10,...,7=more than 60)		5.24
		(1.58)
average grade previous year (0-10)		6.71
		(0.76)
observations	69	58

Note: The table reports the mean of selected variables for the 2012 cohort in 2012 and 2013. Standard deviations are given in parentheses. ECTS are study credits; a student is required to obtain 60 ECTS in order to complete one year of university. ***, **, * indicates a significant difference with the 2012 baseline at the 1%, 5% and 10% level, respectively.

Table 4.3: Summary statistics current members and members who quit the sorority

	current	quit
age	18.48	18.73
	(0.84)	(1.19)
born in The Netherlands	0.93	0.82
	(0.26)	(0.40)
mother born in The Netherlands	0.91	0.82
	(0.28)	(0.40)
father born in The Netherlands	0.95	0.73
	(0.22)	(0.47)
parents married	0.83	0.82
	(0.38)	(0.40)
higher than average household income	0.95	1.00
	(0.22)	(0.00)
mother highly educated	0.74	0.80
	(0.44)	(0.42)
father highly educated	0.89	0.80
	(0.31)	(0.42)
highest high school level	0.79	0.73
	(0.41)	(0.47)
grade final exam	7.07	7.55*
	(0.67)	(1.13)
studies economics	0.14	0.00
	(0.35)	(0.00)
studies business	0.35	0.73**
	(0.48)	(0.47)
studies medicine	0.09	0.09
	(0.29)	(0.30)
follows the highest level of secondary education	0.88	0.91
	(0.33)	(0.30)
happiness	7.71	7.10**
	(0.77)	(1.37)
future happiness (1yr)	8.38	7.82**
	(0.62)	(1.17)
future happiness (5yr)	8.50	8.18
	(0.68)	(0.98)
future happiness (10yr)	8.71	8.55
	(0.79)	(0.93)
past happiness (5yr)	6.91	7.50
	(1.33)	(1.65)
observations	58	11

Note: The table reports the mean of selected variables for the 2012 cohort, split up in a group of current members and a group of members who decided to quit the sorority. Standard deviations are given in parentheses. ***, **, * indicates a significant difference with the 2012 baseline at the 1%, 5% and 10% level, respectively.

Table 4.4: Summary statistics current members and members who quit the sorority [continued]

	current	quit
family important	1.00	1.00
	(0.00)	(0.00)
friends important	1.00	1.00
	(0.00)	(0.00)
religion important	0.04	0.18*
	(0.19)	(0.40)
sports important	0.76	0.73
	(0.43)	(0.47)
party important	0.88	0.90
	(0.33)	(0.32)
politics important	0.33	0.50
	(0.47)	(0.53)
very good or excellent health	0.49	0.36
	(0.50)	(0.50)
does sports at least once a week	0.65	0.73
	(0.48)	(0.47)
weight (kg)	62.33	59.73
	(6.47)	(6.23)
height (cm)	171.38	171.55
	(6.07)	(4.78)
BMI	21.26	20.31
	(2.18)	(2.18)
alcohol at least twice a week	0.95	0.64***
	(0.22)	(0.50)
alcohol at least 5 glasses a night	0.69	0.55
	(0.47)	(0.52)
heterosexual	0.98	1.00
	(0.13)	(0.00)
now in serious relationship	0.30	0.18
	(0.46)	(0.40)
had at least one serious relationship	0.61	0.50
	(0.49)	(0.53)
no bed partners	0.23	0.30
	(0.42)	(0.48)
4 or more bed partners	0.23	0.20
	(0.42)	(0.42)
observations	58	11

Note: The table reports the mean of selected variables for the 2012 cohort, split up in a group of current members and a group of members who decided to quit the sorority. Standard deviations are given in parentheses. ***, **, * indicates a significant difference with the 2012 baseline at the 1%, 5% and 10% level, respectively.

4.3 Empirical methodology

Happiness, behaviour, and study outcomes can be correlated among peers in a social network for three reasons (Fowler and Christakis, 2008). First, *homophily* plays a role if individuals with similar characteristics flock together. Second, *induction* explains the correlation if observed outcomes spread from one individual to another, i.e. it is 'contagious.' The third reason is *confounding*, which occurs when individuals experience the same shocks in their environment. For example, if the ministry of education decides to give all students a 100 euro subsidy, happiness is likely to rise across all sorority members. As a result, happiness is positively correlated but this correlation is merely due to the experience of an identical shock, the subsidy. The challenge of my study is to disentangle homophily, induction, and confounding. The remainder of this section is organised as follows. Subsection 4.3.1 outlines the methodology used to investigate whether there is homophily in club formation. Subection 4.3.2 discusses a model to study induction, while controlling for homophily and confounding.

4.3.1 Homophily

I estimate an 'undirectional dyadic regression' (Fafchamps and Gubert, 2007) in order to study homophily in club formation, which means that I study every possible pair of individuals in a cohort and test whether there is a higher probability that two individuals end up in the same club if they have more similar baseline statistics. Formally, I estimate

$$yc_{ij} = \alpha + \beta_1|X_i - X_j| + \beta_2(Z_i + Z_j) + \beta_3(1_{D_i=D_j}) + \epsilon_{ij}, \tag{4.1}$$

where, $i > j$, yc_{ij} equals 1 if i and j in same club and 0 otherwise, X_i is a matrix for i with variables to be differenced with the respective variables for j (e.g. age and income), Z_i is a matrix for i with variables to be summed with the respective variables for j (e.g. age and income), D_i is a matrix for i with variables that are a match or not with j (e.g. previous city of residence), $1_{D_i=D_j}$ equals 1 if value

for i and j is the same and 0 otherwise, and ϵ_{ij} is the standard error. Standard errors are calculated as in Fafchamps and Gubert (2007).[11] Note that all explanatory variables are measured at $t = 0$, i.e. before club formation. This model only allows for the analysis of selection on observables.

My first hypothesis is that β_1 in Equation 4.2 is positive, i.e. individuals who are less different with respect to the outcome variable have a higher probability of selecting into the same club. Also, I expect that β_3 is positive, i.e. individuals who share a characteristic (e.g. follow the same study major) are more likely to become member of the same club. There is no clear expectation for β_2, as the hypothesis concerning the sum is variable-specific and difficult to predict. For example, relatively old individuals might have more tendency to cluster together ($\beta_2 > 0$), but the same can be argued for young individuals ($\beta_2 < 0$). It should be noted that sufficient degree variation, i.e. variance in the number of connections per individual, is necessary for the identification of the effect of sum variables (Fafchamps and Gubert, 2007).

4.3.2 Induction

I assume that induction is stronger for individuals who are socially closer ('friends') than those who are less close ('non-friends'). In other words, I expect that the outcome of individual j has a stronger impact on individual i if i and j are friends. I can disentangle induction and confounding by testing for the existence of this 'friendship effect.' In case of induction, friends will have an additional impact on each other. In case of confounding, peer effects will be the same for friends and non-friends. After all,

[11] Fafchamps and Gubert (2007):

$$Var(\hat{\beta}) = \frac{1}{N-K}(X'X)^{-1}(\sum_{i=1}^{N}\sum_{j=1}^{N}\sum_{k=1}^{N}\sum_{l=1}^{N}\frac{m_{ijkl}}{2N}X_{ij}u_{ij}u'_{kl}X_{kl})(X'X)^{-1},$$

where β is a vector of coefficients, N is the number of dyadic observations, K is the number of regressors, X is the matrix of all regressors, X_{ij} is the matrix of regressors for dyadic observation ij, m_{ijkl} equals 1 if $i = k$, $j = l$, $i = l$ or $j = k$, and 0 otherwise. The assumptions are that $E[u_{ij}, u_{ik}] \neq 0$, $E[u_{ij}, u_{kj}] \neq 0$, $E[u_{ij}, u_{jk}] \neq 0$, $E[u_{ij}, u_{ki}] \neq 0$ for all k. $E[u_{ij}, u_{km}] = 0$ otherwise. This method also corrects for heteroscedasticity.

everyone is affected by the same shock, regardless of their friendship status.

Formally, I estimate

$$y_{it} = \alpha + \beta_1 y_{it-1} + \beta_2 y_{jt} + \beta_3 y_{jt-1} + \beta_4 f_{ijt} + \beta_5 (f_{ijt} * y_{jt}) + X_i \gamma + \epsilon_{ijt}, \tag{4.2}$$

where y_{it} is outcome measure y for individual i at time t, f_{ijt} equals 1 if i mentions j as friend at time t and 0 otherwise, X_i is a matrix with baseline control variables (age, education, household income), and ϵ_{ijt} is the standard error.[12] I estimate this model for all pairs of clubmates. The null hypothesis of no induction is $\beta_5 = 0$. Depending on whether the dependent variable is continuous or categorical, I estimate this model with ordinary least squared (OLS) or ordered logit, respectively. Standard errors are clustered at the individual level. The inclusion of the lag of the clubmate's outcome variable allows me to control for homophily, assuming that the individuals in the pair did not know each other at $t = 0$.

In order to know whether clubmates identify each other as friends, the follow-up questionnaire asks respondents to identify at most five sorority members whom they are closest to. I identify someone as a 'friend' if she was mentioned as a friend by the respondent. Table 4.5 shows that respondents do not necessarily mention clubmates: on average, respondents mention five names of which three are clubmates.

This methodology is closely related to Fowler and Christakis (2008). It is similar in the sense that we both use relationships to disentangle confounding and induction. Their methodology differs from mine in the sense that I use interaction effects, where Fowler and Christakis (2008) run regressions on different subsets of the data. Several papers have criticised the methodology of Fowler and Christakis (2008) and other related papers by the same authors. Cohen-Cole and Fletcher (2008) argue that friendships can also be based on homophily. However, testing for directionality of friendship, as I will do in the robustness checks, should take care of such concerns. Next to the argument about friendship

[12] I do not include club effects, because these would absorb the effect of the clubmate's outcome variable. The model would then become a model that estimates how outcome variables are distributed around the club mean.

selection, Cohen-Cole and Fletcher (2008) state that the methodology does not include enough contextual factors, does not sufficiently control for selection, and produces biased estimated coefficients.[13] Christakis and Fowler (2013) refute the arguments, referring to sensitivity analyses done by Iwashyna et al. (2011) and thereby confirming the validity of their method.

Table 4.5: Friendship variables

	club1	club2	club3	club4	club5
no. friends in year club	2.923	3.533	3.429	3.100	2.800
	(1.32)	(1.41)	(0.85)	(1.60)	(1.79)
no. friends mentioned	4.923	5.000	4.571	4.400	4.800
	(0.28)	(0.00)	(0.94)	(1.58)	(0.45)
proportion friends in year club	0.600	0.707	0.768	0.711	0.600
	(0.28)	(0.28)	(0.17)	(0.27)	(0.40)
no. club members	13	15	14	10	5

Note: Mean for friendship variables for the 2012 cohort as measured in 2013. Standard deviations are given in parentheses.

4.4 Results

4.4.1 Homophily

Table 4.6 shows the dyadic regression results, i.e. a logit estimation in which the dependent variable equals 1 if a pair of respondents is in the same club and 0 otherwise.[14] Differences in background variables, such as age and household income, as well as in happiness and behavioural variables are statistically insignificant. I find that existing networks matter. Coming from the same city, following

[13] Shalizi and Thomas (2011) also discuss problems with the methodology of Fowler and Christakis (2008).

[14] Note that I only include pairs of members of the same cohort. Therefore, the number of observations is of the magnitude of $(69^2 - 69) + (101^2 - 101) = 14,792$. The actual number of observations is less due to missing observations.

the same major or both living in a sorority house increases the odds of being in the same club by 96%, 29%, and 309%, respectively.[15] The proportion of members living in a sorority house varies considerably across clubs (see Data section), which offers an explanation for the high effect of 309%. I find that the coefficient on the sum of household income is significant at the 5% level. This result implies that respondents are more likely to report to be in the same club if their combined (parental) household income is relatively high. Appendix Table 4.B1 shows that these conclusions also hold for other specifications. In short, I do not find evidence for homophily in happiness, behaviour or educational outcomes. However, parental income and existing networks are important in the sorting process.

4.4.2 Induction

The 2012 cohort is the only cohort for which I have two waves of data. Since the analysis of induction requires at least two data points per individual, I focus on the 2012 cohort in the remainder of this section. For categorical variables, I drop categories with two observations or less. In the following subsection, I test for the robustness against inclusion of these outliers. Table 4.7 shows the estimation of Equation 4.2 for happiness, BMI, sports frequency, alcohol quantity, relationship status, number of bed partners, and average grade,[16] respectively. I find that clubmate's current outcomes are significant for BMI and relationship status. Whereas the effect of clubmate's BMI is positive, the coefficient for relationship status is negative. The interaction effect between the friendship dummy and the clubmate's outcome variable is significantly positive for happiness, sports, alcohol, and bed partners. It is significantly negative for grade. The coefficients of own lagged outcomes are significant and positive, except with regard to happiness and grade.

[15] I use the estimates from model A3. I use that $e^{0.672} = 1.958$, $e^{1.409} = 4.092$ and $e^{0}.258 = 1.294$.

[16] Since the respondents are generally in their first year of university, no baseline university grade is available. I therefore take the final exam grade as the baseline average grade.

Table 4.6: Dyadic regressions pair of individuals in same club (0/1) on selected variables

	A1	A2	A3
constant	-1.617***	-2.578***	-3.163**
	(0.09)	(0.24)	(1.48)
Difference in			
age	0.005	0.014	0.024
	(0.03)	(0.04)	(0.04)
household income (1=very low,...,5=very high)	-0.001	-0.005	-0.019
	(0.07)	(0.07)	(0.07)
happiness (0=very unhappy,...,10=very happy)	-0.075	-0.071	-0.079
	(0.05)	(0.05)	(0.06)
bmi	0.003	-0.005	-0.003
	(0.03)	(0.03)	(0.03)
sports frequency (1=never,...,6=more than 5 times per week)	-0.011	-0.013	-0.014
	(0.04)	(0.04)	(0.04)
alcohol quantity (1=never,...,6=10 or more glasses)	0.013	0.018	0.034
	(0.03)	(0.03)	(0.03)
bed partners (1=0,...,5=10 or more)	-0.030	-0.042	-0.063
	(0.05)	(0.05)	(0.05)
grade final exam (1-10)	-0.002	-0.017	-0.009
	(0.05)	(0.05)	(0.06)
Match in (1=same,0=different)			
previous city of residence		0.669**	0.672**
		(0.30)	(0.29)
lives in a sorority house		1.404***	1.409***
		(0.28)	(0.28)
study major		0.247*	0.258**
		(0.13)	(0.13)
now in serious relationship		0.084	0.080
		(0.10)	(0.10)
highest high school level		0.033	0.033
		(0.12)	(0.12)
Sum of			
age			0.015
			(0.04)
household income (1=very low,...,5=very high)			0.093**
			(0.04)
happiness (0=very unhappy,...,10=very happy)			-0.005
			(0.03)
bmi			-0.009
			(0.01)
sports frequency (1=never,...,6=more than 5 times per week)			-0.022
			(0.02)
alcohol quantity (1=never,...,6=10 or more glasses)			0.019
			(.)
bed partners (1=0,...,5=10 or more)			0.021
			(0.04)
grade final exam (1-10)			-0.022
			(0.04)
obs	11388	11092	11092

Note: Logit regression with dependent variable equals 1 if pair of respondents in same club, 0 otherwise. Sample only includes pairs of respondents from the same cohort. Standard errors, given in parentheses, are calculated using the methodology in Fafchamps and Gubert (2007). *** indicates significance at the 0.01 level; ** indicates significance at the 0.05 level; * indicates significance at the 0.10 level.

Table 4.7: OLS or ordered logistic regression of own current happiness and behaviour on own lagged and clubmate's current and lagged happiness and behaviour

	happiness	bmi	sports	alcohol	relation	bed partners	grade
constant		4.645					
		(4.15)					
own outcome (lag)	0.562	0.933***	1.527***	1.004***	2.298***	3.162***	0.503
	(0.49)	(0.07)	(0.38)	(0.27)	(0.81)	(0.70)	(0.47)
clubmate's outcome (lag)	-0.197	-0.139**	0.126**	0.039	0.198	0.097	0.221*
	(0.12)	(0.05)	(0.05)	(0.07)	(0.25)	(0.11)	(0.12)
clubmate's outcome	-0.102	0.163**	0.041	-0.152	-0.499***	-0.027	0.057
	(0.12)	(0.06)	(0.13)	(0.11)	(0.12)	(0.14)	(0.10)
friend	-4.442***	0.898	-1.032*	-1.548**	0.256	-0.387	3.430***
	(1.59)	(0.86)	(0.60)	(0.76)	(0.21)	(0.32)	(1.24)
friend*clubmate's outcome	0.561***	-0.037	0.256*	0.288*	0.272	0.258*	-0.500***
	(0.20)	(0.04)	(0.14)	(0.16)	(0.43)	(0.14)	(0.19)
own age	-0.357	-0.218	0.526	-0.478	0.544	-0.030	0.577
	(0.43)	(0.22)	(0.37)	(0.32)	(0.52)	(0.42)	(0.36)
own highest high school level	0.802	0.626*	1.559**	-1.100*	1.247	-1.265*	0.506
	(0.72)	(0.36)	(0.68)	(0.64)	(1.43)	(0.71)	(0.70)
own income	0.941	0.047	-1.298***	-0.045	0.892	0.063	1.130*
	(0.67)	(0.29)	(0.47)	(0.47)	(0.78)	(0.58)	(0.66)
Obs	594	624	592	654	650	632	616
R-squared	0.066	0.786	0.217	0.119	0.200	0.335	0.078
Total friend effect=0 (p-value)	0.006	0.023	0.025	0.239	0.511	0.108	0.005
Clustered standard errors	YES	YES	YES	YES	YES	YES	YES
Method	OLOG	OLS	OLOG	OLOG	OLOG	OLOG	OLOG

Note: Dependent variables are own current happiness (0=very unhappy,...,10=very happy), BMI, sports frequency (1=never,...,6=more than 5 times per week), alcohol quantity (1=never,...,6=10 or more glasses), relationship status (1=currently in a relationship, 0=otherwise) and number of bed partners (1=0 bed partners,...,5=10 or more) and average grade (1-10), respectively. Ordered logistic regression results ('OLOG') are reported, except for BMI, for which linear regression results ('OLS') are reported. Standard errors are clustered at the individual level and are given in parentheses. *** indicates significance at the 0.01 level; ** indicates significance at the 0.05 level; * indicates significance at the 0.10 level. For ordered logit estimations, the pseudo R-squared is reported. 'Total friend effect=0 (p-value)' indicates the p-value of a Wald test with null hypothesis *clubmate's outcome + friend*clubmate's outcome=0*.

4.4.3 Robustness

In the previous subsection, I found a significant effect of a clubmate's BMI and relationship status, but no significant additional effect if the clubmate was a friend. Alternatively, I found a significant effect of a friend's outcome variable on own happiness, sports behaviour, alcohol use, bed partners, and average grade. In this subsection, I test different specifications to find out how robust my estimates are. Table 4.8 gives an overview of the robustness checks and focuses on the coefficients of interest: clubmate's outcome variable and the interaction of this variable with the friendship dummy. Panel A shows the original model estimation as was reported in Table 4.7.

In order to disentangle confounding factors from induction, I assume that induction is stronger between friends than between non-friends. In the original model, I identified someone as a 'friend' if she was mentioned as a friend by the respondent. In line with the assumption, I expect that the effect of a mutual friendship is the same or stronger than the estimated effect in the original model. Mutual friendships are likely to be stronger, or at least not less strong, than a friendship that is not reciprocated. Inversely, I expect the effect to be the same or weaker if the respondent mentions a friend while the friend does not reciprocate the friendship (a 'unidirectional friendship from self'). In the original model a friendship is either mutual or uni-directional from self. By definition, the effect for the latter is lower than the originally estimated effect if the effect for mutual friendships is higher. I expect the effect to be smaller or equal if the respondent is mentioned as a friend by the clubmate while she does not mention her back (a 'unidirectional friendship from clubmate'). The reason for this conjecture follows from the original assumption: an individual is more affected by someone she identifies as a friend. Using the directionality of friendships as a robustness check is in line with Fowler and Christakis (2008).

Panels B, C, and D show the effect of a unidirectional friendship from clubmate, a uni-directional friendship from self, and a mutual friendship, respectively (see Tables 4.C1, 4.C2 and 4.C3 for full results). In case of a unidirectional friendship from the clubmate, the coefficient on the interaction effect between the friendship variable and the outcome variable ceases to be significant. On the other

hand, effects of mutual friendships are significant and higher for happiness, alcohol use, and average grade. The interaction effects for these variables are not significant when looking at the unidirectional relationship from self, leading to the conclusion that the original results were mainly driven by mutual friendships. The coefficients on the friendship interaction for sports behaviour and bed partners are insignificant for mutual friendships. Hence, though my assumption seems to hold for other variables, results for sports and bed partners are not robust.

As there are few missing observations for the outcome variables,[17] I simply dropped observations with missing values. I now impute missing values for outcome variables with the median and include a dummy for missing observations in the model (see Table 4.C4 for full results). Panel E shows that the original results are robust to the imputation of missing values.

In the original model, I drop categories for which there were few observations.[18] Panel F reports the estimations if I include outliers. Estimated results are generally the same as in the original model, with one notable exception: the coefficient on the interaction between friendship and happiness more than halves and ceases to be significant.

I assumed that members only meet at their entry into the sorority. In the survey, I included a question that asked the respondent to list members from her cohort whom she knew from before her membership (a 'previous acquaintance'). Among all pairs of clubmates, 6.5% were previous acquaintances. Panel G shows the results if I exclude these pairs. Coefficient estimates are qualitatively the same, except for sports behaviour.

To conclude, results with regards to BMI, alcohol use, relationship status, and average grade pass do not qualitatively change across the robustness checks in this subsection. Happiness is generally robust, but the estimated effect disappears when outliers are included. The results for sports frequency and bed

[17] In the latest wave, there are zero missing values for sports frequency, alcohol quantity, bed partners, and relationship status; one for happiness; and two for average grade and BMI.

[18] I drop categories with two observations or less. This implies that I drop category 4 (1 observation), 5 (1) and 6 (1) for happiness; 1 (2) and 6 (1) for sports frequency; and 2 (1) for alcohol quantity. I drop no categories for relationship status, bed partners, and average grade.

partners are not robust. Besides the robustness checks discussed in this subsection, Appendix Tables 4.C7-4.C13 show results for different sets of regressors and Appendix Tables 4.C14-4.C19 show non-linear specifications. These tables show similar estimates to the original model and therefore confirm the conclusions regarding robustness in this subsection.

4.4.4 Interpretation

In this subsection, I will discuss the interpretation of the estimated coefficients. Since the previous subsection discarded the robustness of the results for sports frequency and bed partners, I will exclude these outcome variables from further discussion.

Although clubmate's BMI and relationship status are significant, there is no additional friendship effect. These outcome variables seem therefore to be subject to confounding factors rather than induction. A one standard-deviation increase in clubmate's BMI relates to an increase in own BMI of 0.358, which amounts to gaining one kilo at the average height.[19] The coefficient of relationship status implies that one is 6.5% *less* likely to be in a relationship if one's clubmate is.

I find a significant and robust friendship effect for happiness, alcohol, and grade. I therefore conclude that these outcome variables are subject to induction. To get a better understanding of the effects, I calculate the difference in predicted probability between friends and non-friends at different levels of the clubmate's outcome variable. Figure 4.2 displays this difference for different levels of own happiness. The graphs show that an individual is more likely to report a low (high) level of happiness if a befriended clubmate has a low (high) level, relative to a clubmate with a similar level of happiness who is not a friend. For example, an individual is 7% more likely (p-value 0.090) to report that her happiness equals seven if a befriended clubmate also reported seven, while she is 8% less likely to report a nine (p-value 0.039).

[19]The standard deviation of BMI is 2.196. The average height is 1.714. The formula for BMI is $\frac{weight(kg)}{(height(m))^2}$. Hence an increase of 0.358 is associated with a $1.714^2 * 0.358 = 1.051$ kilo increase.

Table 4.8: Robustness checks

	happiness	bmi	sports	alcohol	relation	bed partners	grade
Panel A: Original (Table 4.7)							
clubmate's outcome	-0.102	0.163**	0.041	-0.152	-0.499***	-0.027	0.057
	(0.12)	(0.06)	(0.13)	(0.11)	(0.12)	(0.14)	(0.10)
friend*clubmate's outcome	0.561***	-0.037	0.256*	0.288*	0.272	0.258*	-0.500***
	(0.20)	(0.04)	(0.14)	(0.16)	(0.43)	(0.14)	(0.19)
Panel B: Unidirectional friendship from clubmate (Table 4.C1)							
clubmate's outcome	0.063	0.153**	0.083	-0.088	-0.333***	0.030	-0.052
	(0.11)	(0.06)	(0.11)	(0.08)	(0.09)	(0.13)	(0.09)
unidirectional*clubmate's outcome	-0.238	0.028	0.440	0.372	-0.556	0.274	-0.353
	(0.37)	(0.05)	(0.54)	(0.35)	(0.71)	(0.35)	(0.32)
Panel C: Unidirectional friendship from self (Table 4.C2)							
clubmate's outcome	0.044	0.159**	0.050	-0.067	-0.457***	0.001	-0.049
	(0.11)	(0.06)	(0.12)	(0.09)	(0.10)	(0.14)	(0.08)
unidirectional*clubmate's outcome	-0.055	-0.039	0.480**	-0.043	0.682	0.438**	-0.280
	(0.34)	(0.06)	(0.23)	(0.23)	(0.49)	(0.21)	(0.23)
Panel D: Mutual friendship (Table 4.C3)							
clubmate's outcome	-0.091	0.160**	0.116	-0.145	-0.485***	0.028	0.002
	(0.11)	(0.06)	(0.13)	(0.10)	(0.10)	(0.14)	(0.09)
mutual*clubmate's outcome	0.820***	-0.030	-0.025	0.424**	0.219	0.061	-0.522**
	(0.26)	(0.06)	(0.24)	(0.20)	(0.59)	(0.15)	(0.22)
Panel E: Missing observations imputed (Table 4.C4)							
clubmate's outcome	-0.105	0.170**	0.054	-0.152	-0.498***	-0.035	0.099
	(0.12)	(0.06)	(0.13)	(0.11)	(0.12)	(0.13)	(0.10)
friend*clubmate's outcome	0.553***	-0.043	0.272*	0.288*	0.250	0.251*	-0.558***
	(0.20)	(0.04)	(0.14)	(0.16)	(0.43)	(0.14)	(0.19)
Panel F: Outliers included (Table 4.C5)							
clubmate's outcome	0.010	0.163**	-0.055	-0.176	-0.499***	-0.027	0.057
	(0.10)	(0.06)	(0.12)	(0.11)	(0.12)	(0.14)	(0.10)
friend*clubmate's outcome	0.211	-0.037	0.251**	0.272*	0.272	0.258*	-0.500***
	(0.21)	(0.04)	(0.12)	(0.16)	(0.43)	(0.14)	(0.19)
Panel G: Previous acquaintances excluded (Table 4.C6)							
clubmate's outcome	-0.074	0.167**	0.022	-0.132	-0.474***	-0.021	0.037
	(0.12)	(0.06)	(0.13)	(0.11)	(0.16)	(0.14)	(0.10)
friend*clubmate's outcome	0.537***	-0.025	0.173	0.302*	0.569	0.242*	-0.623***
	(0.20)	(0.04)	(0.14)	(0.17)	(0.45)	(0.14)	(0.20)
Clustered standard errors	YES	YES	YES	YES	YES	YES	YES
Method	OLOG	OLS	OLOG	OLOG	OLOG	OLOG	OLOG

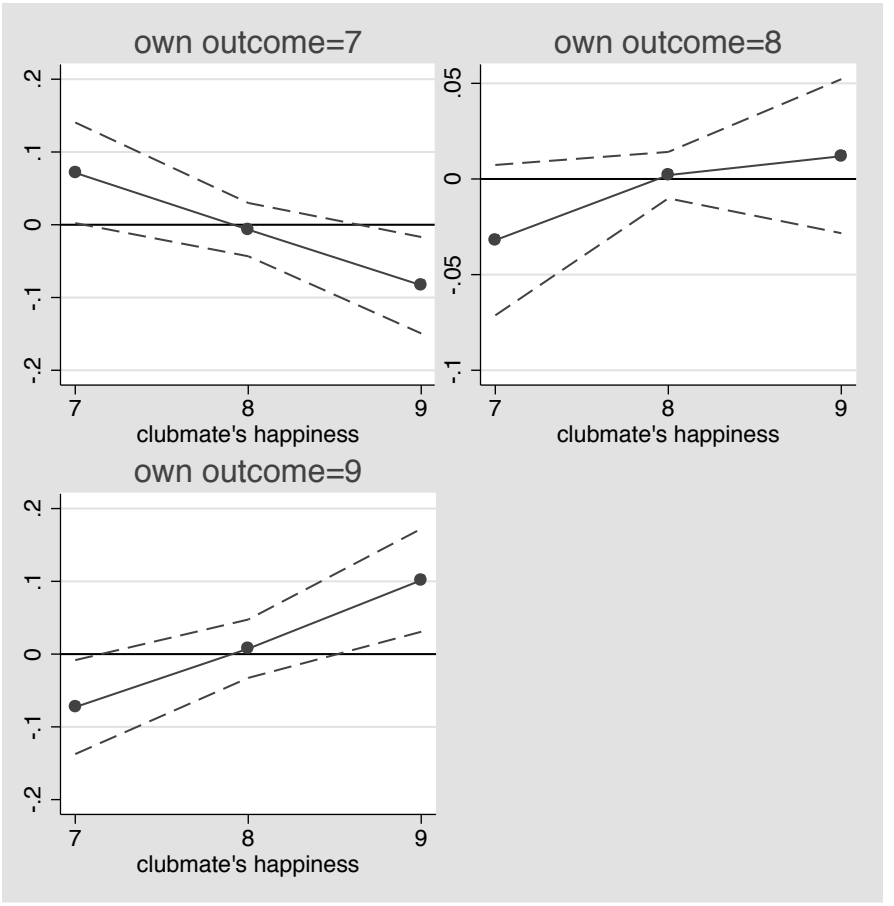
Note: Table displays results of five robustness checks. Only the main coefficients of interest are reported; the full results can be found in the table referred to in brackets. Dependent variables are own current happiness (0=very unhappy,...,10=very happy), BMI, sports frequency (1=never,...,6=more than 5 times per week), alcohol quantity (1=never,...,6=10 or more glasses), relationship status (1=currently in a relationship, 0=otherwise) and number of bed partners (1=0 bed partners,...,5=10 or more) and average grade (1-10), respectively. Ordered logistic regression results ('OLOG') are reported, except for BMI, for which linear regression results ('OLS') are reported. Standard errors are clustered at the individual level and are given in parentheses. *** indicates significance at the 0.01 level; ** indicates significance at the 0.05 level; * indicates significance at the 0.10 level.

Figure 4.3 shows a similar relationship for alcohol quantity: a respondent is more likely to report low (high) levels of alcohol use if a befriended clubmate also reports low (high) levels, relative to a clubmate who is not reported to be a friend. The difference in probabilities is mainly significant at low levels of clubmate's outcome for alcohol use. For example, a respondent is 11% less likely (p-value 0.034) to report the highest drinking frequency (more than 10 glasses) if a befriended clubmate reported the lowest category (3-4 glasses).

Last, Figure 4.4 shows that there is an inverse relationship between own grade and the effect of a friend's grade. If a befriended clubmate's grade is low, an individual is more likely to have a high grade than if a clubmate is not a friend. The difference in probability is only significant at the 10% level for low values of clubmate's average grade. For example, one is 5% less likely to report a five (p-value 0.096) if a befriended clubmate also reported a five, while one is 10% more likely (p-value 0.039) to report an eight if a befriended clubmate reported a five. I will discuss possible explanations for this effect in the conclusion.

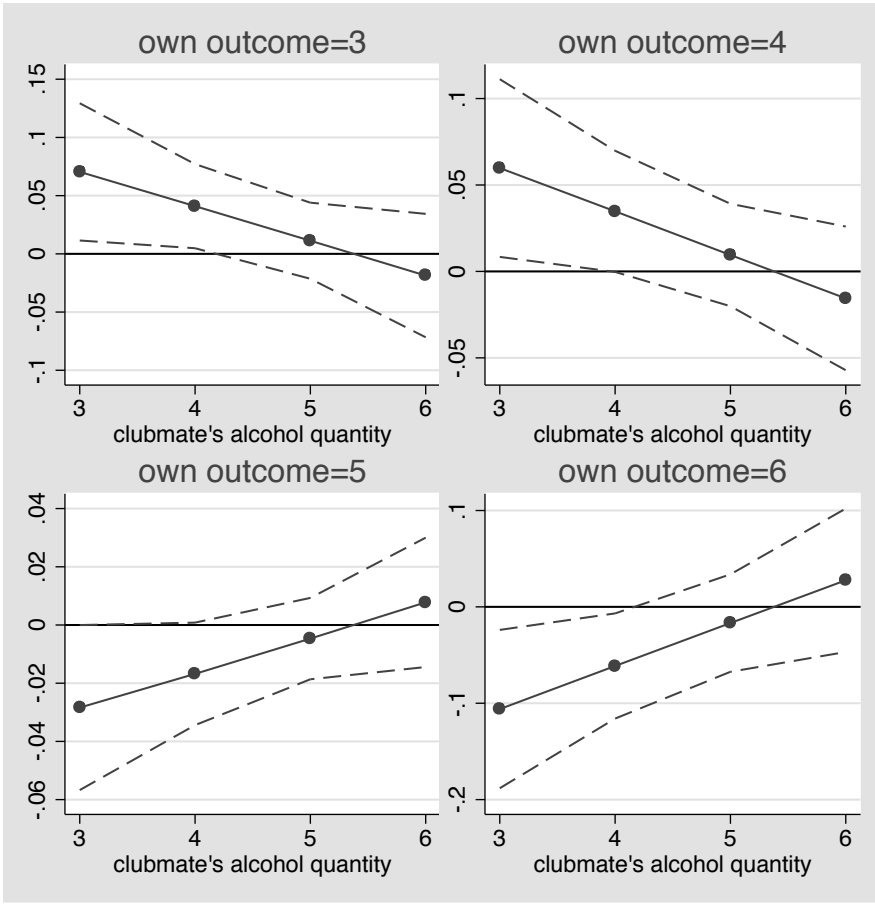
To summarise, BMI and relationship status are correlated between clubmates. As I find no significant friendship effect and the dyadic regression showed no signs of homophily, this correlation should probably be explained by confounding factors. In the next section, I will further elaborate on what such factors could be. Furthermore, one is more likely to report happiness and alcohol use similar to that of a clubmate if the clubmate is reported to be a friend. The inverse is true for grades: if a clubmate is a friend, one is more likely to report a grade that is opposite to the clubmate's grade. Because the impact of a befriended clubmate is stronger than that of a regular clubmate, I conclude that these outcome variables are likely subject to induction. However, robustness checks show that these effects are driven by mutual friendships. Results should therefore be interpreted with caution, which I will elaborate on in the following section.

Figure 4.2: Difference in predicted probability for happiness



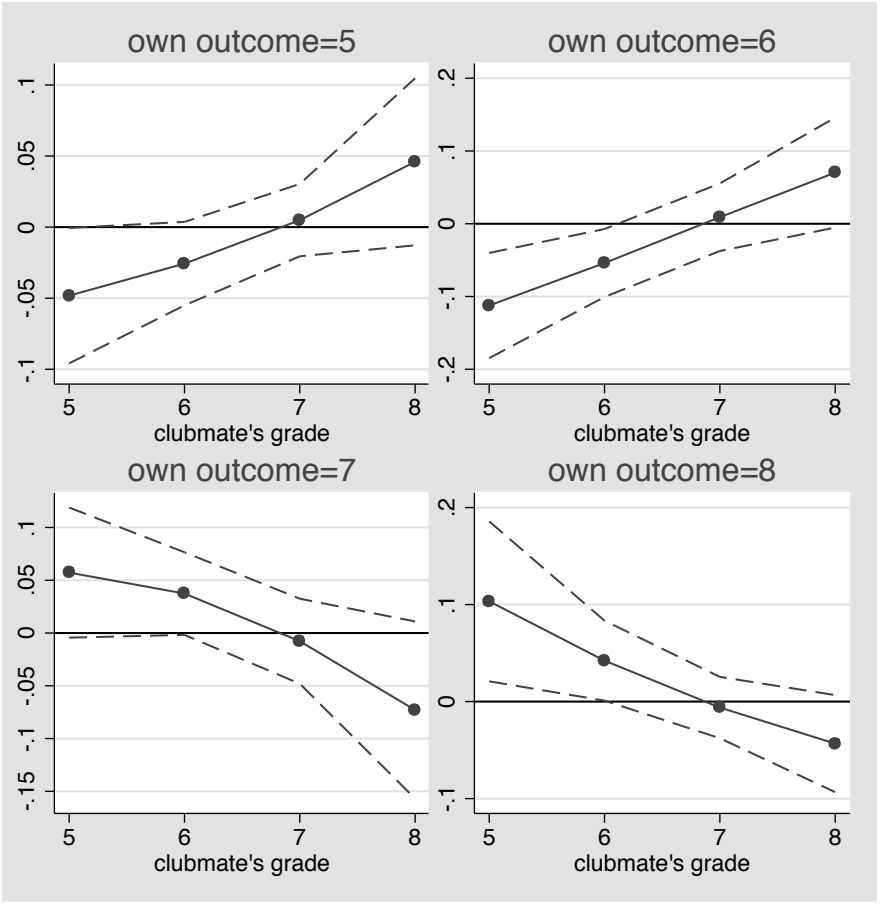
Note: Figure displays the difference in predicted probability between clubmates that are friends and those that are not. Probabilities are plotted at different levels of clubmate's happiness for own happiness of 7, 8 and 9. Remaining variables are taken at average value. Estimates are based on Table 4.7. Dashed lines indicate 10% confidence interval.

Figure 4.3: Difference in predicted probability for alcohol quantity



Note: Figure displays the difference in predicted probability between clubmates that are friends and those that are not. Probabilities are plotted at different levels of clubmate's alcohol quantity for categories 3 (3-4 glasses on a drinking night), 4 (5-6 glasses), 5 (7-9 glasses), and 6 (10 or more glasses). Remaining variables are taken at average value. Estimates are based on Table 4.7. Dashed lines indicate 10% confidence interval.

Figure 4.4: Difference in predicted probability for average grade



Note: Figure displays the difference in predicted probability between clubmates that are friends and those that are not. Probabilities are plotted at different levels of clubmate's average grade for own average grade of 5, 6, 7 and 8. Remaining variables are taken at average value. Estimates are based on Table 4.7. Dashed lines indicate 10% confidence interval.

4.5 Discussion and conclusion

This paper discusses the spread of happiness, behaviour, and educational outcomes in a social network. Using unique data from a Dutch sorority, I study a network in which social ties exist between fellow members of 'clubs.' I do not find that respondents with similar happiness, behaviour or educational outcomes cluster together. However, parental income and existing networks are important in the club sorting process. Moreover, I find that correlations of BMI and relationship status are due to confounding factors. Members all experience the same 'shock' of being in the first year of college, which is likely to affect their BMI. As a result, BMI is positively correlated among clubmates. The negative correlation in relationship status can be explained by the fact that some members became single or entered a relationship during the year. These changes are fairly balanced across clubs, causing the correlation to be negative.

Furthermore, happiness, alcohol use, and grades are likely to spread through a social network. My findings with regards to happiness are in line with Fowler and Christakis (2008) who also find positive effects of happiness. In addition, the positive effect of a peer's alcohol consumption is in line with findings in Reifman et al. (2006). The negative effect of a peer's grade, however, is in contrast with the positive effect found in Sacerdote (2001). This difference in findings can be explained by the fact that Sacerdote (2001) uses random allocation of peers, whereas peers select each other in my study. Moreover, I look at a sorority setting, while Sacerdote (2001) focuses on a more heterogeneous student population.

Robustness checks reveal that the negative effect of a peer's grade is robust. This negative effect could be explained in two ways. First, respondents could engage in compensatory behaviours, offsetting the other's good (bad) grades with bad (good) grades. Second, this finding could be a consequence of the fact that different individuals seek each other out as friends. I find no significant effect of difference in grades in the dyadic regression, which implies that grades do not play a role in the club sorting process. However, grade differences could play a role in forming friendships within the club. I find

that the negative effect is mainly driven by mutual friendships, whereas there is no significant effect for unidirectional friendships. Hence, the direction of the effect is not clear and it might be explained by selection effects in friendships.

These remarks apply more generally to my findings: significant induction effects are driven by mutual friendships. I am able to control for homophily effects at the club level - by including clubmate's lagged outcome variable - and confounding factors at the club level - by using friendship effects. However, I do not control for homophily in friendships. Hence, the significant effects I find can be either explained by induction or by homophily in friendships. Future studies could collect data more frequently in order to obtain a better understanding of the evolution of friendships.

Christakis and Fowler (2013) indicate the value of generative models (Steglich et al., 2006, 2010), which model the coevolution of networks and behaviour and thereby handle the endogeneity problems in studying peer effects. Future research on happiness and behaviour in social networks could use such generative models to get a more thorough understanding of the dynamics of a social network, including friendship formation. Moreover, the number of observations in this study is small. It would therefore be good to use larger data sets in order to test the robustness of the findings. Also, three data waves (or more) would enable researchers to control for unobservable heterogeneity using panel data models.

Policymakers would benefit from knowing how happiness, unhealthy behaviour or academic performance is affected by peers. Proper policy design depends on the channels behind peer effects. Outcomes that spread through a social network can be influenced efficiently by focusing on central individuals in a social network. For example, a campaign against binge drinking could target a few popular individuals in the sorority. A change in behaviour by such girls would affect their peers. Such an approach would not work if behaviour is a result of common shocks to the environment or homophily. Campaigns should then be directed at all members of the target group. Also, policy could look further into the root causes of the common shocks in order to learn whether they can be tackled directly.

4.A Summary statistics

Table 4.A1: Summary statistics baseline variables 2012 cohort

Variable	Mean	Std. Dev.	N
age	18.522	0.901	69
born in The Netherlands	0.913	0.284	69
mother born in The Netherlands	0.899	0.304	69
father born in The Netherlands	0.913	0.284	69
parents married	0.826	0.382	69
higher than average household income	0.956	0.207	68
mother highly educated	0.75	0.436	68
father highly educated	0.879	0.329	66
highest high school level	0.783	0.415	69
grade final exam (1-10)	7.145	0.772	69
studies economics	0.118	0.325	68
studies business	0.412	0.496	68
studies medicine	0.088	0.286	68
follows the highest level of secondary education	0.882	0.325	68
lives in a sorority house	0.478	0.503	69
happiness	7.618	0.898	68
future happiness (1yr)	8.290	0.75	69
future happiness (5yr)	8.449	0.738	69
future happiness (10yr)	8.681	0.813	69
past happiness (5yr)	7	1.382	68
family important	1	0	69
friends important	1	0	69
religion important	0.059	0.237	68
sports important	0.754	0.434	69
party important	0.882	0.325	68
politics important	0.353	0.481	68
very good or excellent health	0.471	0.503	68
does sports at least once a week	0.662	0.477	68
weight (kg)	61.91	6.459	68
height (cm)	171.406	5.852	69
BMI	21.105	2.196	68
alcohol at least twice a week	0.899	0.304	69
alcohol at least 5 glasses a night	0.667	0.475	69
heterosexual	0.986	0.12	69
now in serious relationship	0.279	0.452	68
had at least one serious relationship	0.597	0.494	67
no bed partners	0.239	0.43	67
4 or more bed partners	0.224	0.42	67

Table 4.A2: Summary statistics baseline variables 2013 cohort

Variable	Mean	Std. Dev.	N
age	18.574	0.864	101
born in The Netherlands	0.921	0.271	101
mother born in The Netherlands	0.881	0.325	101
father born in The Netherlands	0.9	0.302	100
parents married	0.733	0.445	101
higher than average household income	0.891	0.313	101
mother highly educated	0.747	0.437	99
father highly educated	0.818	0.388	99
highest high school level	0.822	0.385	101
grade final exam (1-10)	7.101	0.707	99
studies economics	0.04	0.197	100
studies business	0.28	0.451	100
studies medicine	0.1	0.302	100
follows the highest level of secondary education	0.871	0.337	101
lives in a sorority house	0.386	0.489	101
happiness	7.267	0.968	101
future happiness (1yr)	8.040	0.747	101
future happiness (5yr)	8.297	0.769	101
future happiness (10yr)	8.406	1.002	101
past happiness (5yr)	7.178	1.438	101
family important	1	0	100
friends important	1	0	100
religion important	0.06	0.239	100
sports important	0.810	0.394	100
party important	0.9	0.302	100
politics important	0.39	0.49	100
very good or excellent health	0.45	0.5	100
does sports at least once a week	0.722	0.451	97
weight (kg)	61.197	6.696	99
height (cm)	170.38	6.53	100
BMI	21.119	2.056	99
alcohol at least twice a week	0.909	0.289	99
alcohol at least 5 glasses a night	0.778	0.418	99
heterosexual	1	0	100
now in serious relationship	0.27	0.446	100
had at least one serious relationship	0.690	0.465	100
no bed partners	0.143	0.352	91
4 or more bed partners	0.242	0.431	91

Table 4.A3: Selected variables as measured in 2012 for 2012 clubs

	club1	club2	club3	club4	club5
age	18.154	18.467	18.571	18.538	19.083
	(0.69)	(0.83)	(0.65)	(0.52)	(1.44)
hh relative income	4.385	4.200	4.214	4.231	4.273
	(0.51)	(0.68)	(0.43)	(0.60)	(0.47)
highest high school level	0.615	0.867	0.857	0.923	0.583
	(0.51)	(0.35)	(0.36)	(0.28)	(0.51)
happiness	7.692	7.867	7.500	7.846	7.273
	(0.85)	(0.99)	(0.65)	(0.55)	(1.10)
bmi	21.601	20.432	21.472	21.707	20.182
	(1.72)	(2.28)	(2.65)	(2.35)	(1.73)
sports frequency	3.615	3.714	3.714	3.308	4.000
	(1.12)	(0.47)	(1.33)	(1.18)	(0.74)
alcohol quantity	4.077	4.000	4.071	4.308	3.583
	(0.95)	(0.76)	(1.07)	(1.18)	(1.24)
now in serious relationship	0.462	0.467	0.143	0.077	0.273
	(0.52)	(0.52)	(0.36)	(0.28)	(0.47)
bed partners	1.077	0.800	1.231	1.538	1.455
	(0.76)	(0.41)	(1.01)	(1.45)	(1.37)
follows the highest level of secondary education	0.833	0.933	0.929	0.846	0.833
	(0.39)	(0.26)	(0.27)	(0.38)	(0.39)
studies economics	0.077	0.267	0.143	0.000	0.083
	(0.28)	(0.46)	(0.36)	(0.00)	(0.29)
studies business	0.231	0.267	0.429	0.667	0.417
	(0.44)	(0.46)	(0.51)	(0.49)	(0.51)
studies medicine	0.077	0.133	0.143	0.000	0.083
	(0.28)	(0.35)	(0.36)	(0.00)	(0.29)
no. club members	13	15	14	13	12

Note: This table reports summary statistics by club for selected variables measured in 2012 for the 2012 cohort. Statistics given are mean and standard deviation (in brackets).

Table 4.A4: Selected variables as measured in 2013 for 2013 clubs

	club1	club2	club3	club4	club5	club6	club7
age	18.667	18.600	18.500	18.733	18.857	18.500	18.231
	(0.98)	(1.06)	(0.65)	(0.70)	(0.77)	(1.02)	(0.73)
hh relative income	4.333	4.067	4.143	3.867	4.429	4.000	3.923
	(0.49)	(0.70)	(0.36)	(0.83)	(0.65)	(0.55)	(0.64)
highest high school level	0.933	0.667	0.929	0.800	0.786	0.857	0.769
	(0.26)	(0.49)	(0.27)	(0.41)	(0.43)	(0.36)	(0.44)
happiness	7.400	7.000	7.643	7.333	6.857	7.000	7.692
	(0.99)	(0.93)	(0.74)	(0.62)	(0.66)	(1.57)	(0.85)
bmi	20.580	21.088	21.167	20.627	22.065	21.099	21.409
	(1.15)	(1.81)	(2.41)	(1.24)	(2.40)	(1.66)	(3.19)
sports frequency	4.000	3.933	3.857	3.600	4.214	3.923	3.615
	(0.85)	(0.80)	(0.95)	(1.18)	(0.97)	(0.86)	(1.19)
alcohol quantity	4.133	4.400	3.929	4.429	4.143	3.538	3.923
	(1.06)	(0.74)	(0.83)	(1.02)	(1.10)	(1.45)	(0.76)
now in serious relationship	0.333	0.267	0.357	0.400	0.143	0.077	0.231
	(0.49)	(0.46)	(0.50)	(0.51)	(0.36)	(0.28)	(0.44)
bed partners	1.400	1.000	1.167	1.643	1.071	0.833	1.111
	(0.83)	(0.55)	(0.58)	(0.93)	(0.92)	(0.83)	(0.93)
follows the highest level of secondary education	0.933	0.733	1.000	0.867	0.857	0.929	0.769
	(0.26)	(0.46)	(0.00)	(0.35)	(0.36)	(0.27)	(0.44)
studies economics	0.067	0.000	0.143	0.000	0.000	0.071	0.000
	(0.26)	(0.00)	(0.36)	(0.00)	(0.00)	(0.27)	(0.00)
studies business	0.333	0.214	0.357	0.400	0.071	0.357	0.231
	(0.49)	(0.43)	(0.50)	(0.51)	(0.27)	(0.50)	(0.44)
studies medicine	0.267	0.000	0.071	0.000	0.143	0.214	0.000
	(0.46)	(0.00)	(0.27)	(0.00)	(0.36)	(0.43)	(0.00)
no. club members	15	15	14	15	14	14	13

Note: This table reports summary statistics by club for selected variables measured in 2013 for the 2013 cohort. Statistics given are mean and standard deviation (in brackets).

4.B Homophily

Table 4.B1: Dyadic regressions pair of individuals in same club (0/1) on selected variables, different estimations

	A1	A2	A3	B1	B2	B3	C1	C2	C3
constant	-1.617***	-2.578***	-3.163**	0.165***	0.072***	-0.001	-3.163***	-3.163***	-3.163***
	(0.09)	(0.24)	(1.48)	(0.01)	(0.02)	(0.17)	(1.18)	(1.18)	(1.18)
Difference in									
age	0.005	0.014	0.024	0.001	0.002	0.003	0.024	0.024	0.024
	(0.03)	(0.04)	(0.04)	(0.00)	(0.01)	(0.00)	(0.03)	(0.03)	(0.03)
household income (1=very low,...,5=very high)	-0.001	-0.005	-0.019	-0.000	-0.001	-0.002	-0.019	-0.019	-0.019
	(0.07)	(0.07)	(0.07)	(0.01)	(0.01)	(0.01)	(0.05)	(0.05)	(0.05)
happiness (0=very unhappy,...,10=very happy)	-0.075	-0.071	-0.079	-0.009*	-0.008	-0.009	-0.079**	-0.079**	-0.079**
	(0.05)	(0.05)	(0.06)	(0.01)	(0.01)	(0.01)	(0.04)	(0.04)	(0.04)
bmi	0.003	-0.005	-0.003	0.000	-0.001	-0.001	-0.003	-0.003	-0.003
	(0.03)	(0.03)	(0.03)	(0.00)	(0.00)	(0.00)	(0.02)	(0.02)	(0.02)
sports frequency (1=never,...,6=more than 5 times per week)	-0.011	-0.013	-0.014	-0.001	-0.002	-0.002	-0.014	-0.014	-0.014
	(0.04)	(0.04)	(0.04)	(0.00)	(0.00)	(0.01)	(0.03)	(0.03)	(0.03)
alcohol quantity (1=never,...,6=10 or more glasses)	0.013	0.018	0.034	0.002	0.002	0.004	0.034	0.034	0.034
	(0.03)	(0.03)	(0.03)	(0.00)	(0.00)	(0.00)	(0.03)	(0.03)	(0.03)
bed partners (1=0,...,5=10 or more)	-0.030	-0.042	-0.063	-0.004	-0.005	-0.008	-0.063*	-0.063*	-0.063*
	(0.05)	(0.05)	(0.05)	(0.01)	(0.01)	(0.01)	(0.04)	(0.04)	(0.04)
grade final exam (1-10)	-0.002	-0.017	-0.009	-0.000	-0.002	-0.001	-0.009	-0.009	-0.009
	(0.05)	(0.05)	(0.06)	(0.01)	(0.01)	(0.01)	(0.04)	(0.04)	(0.04)
Match in (1=same,0=different)									
previous city of residence		0.669**	0.672**		0.099*	0.099*	0.672***	0.672***	0.672***
		(0.30)	(0.29)		(0.05)	(0.05)	(0.19)	(0.19)	(0.19)
house		1.404***	1.409***		0.165***	0.165***	1.409***	1.409***	1.409***
		(0.28)	(0.28)		(0.03)	(0.03)	(0.15)	(0.15)	(0.15)
major		0.247*	0.258**		0.032*	0.034**	0.258***	0.258***	0.258***
		(0.13)	(0.13)		(0.02)	(0.02)	(0.09)	(0.09)	(0.09)
now in serious relationship		0.084	0.080		0.010	0.010	0.080	0.080	0.080
		(0.10)	(0.10)		(0.01)	(0.01)	(0.07)	(0.07)	(0.07)
highest high school level		0.033	0.033		0.004	0.004	0.033	0.033	0.033
		(0.12)	(0.12)		(0.01)	(0.01)	(0.08)	(0.08)	(0.08)
Sum of									
age			0.015			0.002	0.015	0.015	0.015
			(0.04)			(0.00)	(0.03)	(0.03)	(0.03)
household income (1=very low,...,5=very high)			0.093**			0.011**	0.093***	0.093***	0.093***
			(0.04)			(0.01)	(0.03)	(0.03)	(0.03)
happiness (0=very unhappy,...,10=very happy)			-0.005			-0.000	-0.005	-0.005	-0.005
			(0.03)			(0.00)	(0.02)	(0.02)	(0.02)
bmi			-0.009			-0.001	-0.009	-0.009	-0.009
			(0.01)			(0.00)	(0.01)	(0.01)	(0.01)
sports frequency (1=never,...,6=more than 5 times per week)			-0.022			-0.003	-0.022	-0.022	-0.022
			(0.02)			(0.00)	(0.02)	(0.02)	(0.02)
alcohol quantity (1=never,...,6=10 or more glasses)			0.019			0.002	0.019	0.019	0.019
			(.)			(.)	(0.01)	(0.01)	(0.01)
bed partners (1=0,...,5=10 or more)			0.021			0.003	0.021	0.021	0.021
			(0.04)			(0.00)	(0.03)	(0.03)	(0.03)
grade final exam (1-10)			-0.022			-0.003	-0.022	-0.022	-0.022
			(0.04)			(0.01)	(0.03)	(0.03)	(0.03)
obs	11388	11092	11092	11388	11092	11092	11092	11092	11092

Note: Dependent variable equals 1 if pair of individuals is in same club, 0 otherwise. Sample only includes pairs of respondents from the same cohort. Different specifications are (A) logistic regression with standard errors clustered at the individual level, (B) OLS with standard errors as in Fafchamps and Gubert (2007) and (C) logistic regression with standard errors as in Fafchamps and Gubert (2007). Standard errors are given in parentheses. *** indicates significance at the 0.01 level; ** indicates significance at the 0.05 level; * indicates significance at the 0.10 level.

4.C Robustness

4.C.1 Other specifications

Table 4.C1: OLS or ordered logistic regression of own current happiness and behaviour on own lagged and clubmate's current and lagged happiness and behaviour (unidirectional friendship from clubmate)

	happiness	bmi	sports	alcohol	relation	bed partners	grade
constant		4.945					
		(4.22)					
own outcome (lag)	0.537	0.932***	1.522***	1.007***	2.297***	3.175***	0.507
	(0.50)	(0.06)	(0.38)	(0.27)	(0.82)	(0.70)	(0.47)
clubmate's outcome (lag)	-0.189	-0.141**	0.122***	0.028	0.144	0.111	0.226**
	(0.12)	(0.06)	(0.05)	(0.07)	(0.26)	(0.11)	(0.11)
clubmate's outcome	0.063	0.153**	0.083	-0.088	-0.333***	0.030	-0.052
	(0.11)	(0.06)	(0.11)	(0.08)	(0.09)	(0.13)	(0.09)
unidirectional	1.808	-0.699	-1.552	-1.429	-0.587	-0.184	2.621
	(2.93)	(1.02)	(2.06)	(1.54)	(0.57)	(0.50)	(2.12)
unidirectional*clubmate's outcome	-0.238	0.028	0.440	0.372	-0.556	0.274	-0.353
	(0.37)	(0.05)	(0.54)	(0.35)	(0.71)	(0.35)	(0.32)
own age	-0.379	-0.218	0.529	-0.494	0.574	-0.036	0.582
	(0.42)	(0.22)	(0.37)	(0.31)	(0.52)	(0.42)	(0.36)
own highest high school level	0.851	0.634*	1.533**	-1.101*	1.278	-1.255*	0.489
	(0.74)	(0.36)	(0.68)	(0.64)	(1.44)	(0.72)	(0.69)
own income	0.946	0.046	-1.289***	-0.031	0.878	0.063	1.125*
	(0.66)	(0.29)	(0.47)	(0.47)	(0.79)	(0.58)	(0.66)
Obs	594	624	592	654	650	632	616
R-squared	0.063	0.785	0.217	0.118	0.204	0.335	0.076
Total friend effect=0 (p-value)	0.612	0.017	0.342	0.432	0.179	0.392	0.155
Clustered standard errors	YES	YES	YES	YES	YES	YES	YES
Method	OLOG	OLS	OLOG	OLOG	OLOG	OLOG	OLOG

Note: Estimations for unidirectional friendship from clubmate. Dependent variables are own current happiness (0=very unhappy,...,10=very happy), BMI, sports frequency (1=never,...,6=more than 5 times per week), alcohol quantity (1=never,...,6=10 or more glasses), relationship status (1=currently in a relationship, 0=otherwise) and number of bed partners (1=0 bed partners,...,5=10 or more) and average grade (1-10), respectively. Ordered logistic regression results ('OLOG') are reported, except for BMI, for which linear regression results ('OLS') are reported. Standard errors are clustered at the individual level and are given in parentheses. *** indicates significance at the 0.01 level; ** indicates significance at the 0.05 level; * indicates significance at the 0.10 level. For ordered logit estimations, the pseudo R-squared is reported. 'Total friend effect=0 (p-value)' indicates the p-value of a Wald test with null hypothesis *clubmate's outcome + friend*clubmate's outcome=0*.

Table 4.C2: OLS or ordered logistic regression of own current happiness and behaviour on own lagged and clubmate's current and lagged happiness and behaviour (unidirectional friendship from self)

	happiness	bmi	sports	alcohol	relation	bed partners	grade
constant		4.805					
		(4.19)					
own outcome (lag)	0.544	0.934***	1.529***	1.006***	2.339***	3.177***	0.512
	(0.49)	(0.07)	(0.38)	(0.28)	(0.81)	(0.70)	(0.47)
clubmate's outcome (lag)	-0.192	-0.141**	0.124***	0.034	0.184	0.102	0.218*
	(0.12)	(0.05)	(0.05)	(0.07)	(0.25)	(0.12)	(0.12)
clubmate's outcome	0.044	0.159**	0.050	-0.067	-0.457***	0.001	-0.049
	(0.11)	(0.06)	(0.12)	(0.09)	(0.10)	(0.14)	(0.08)
unidirectional	0.436	1.052	-1.822*	-0.078	0.788**	-0.924*	1.813
	(2.81)	(1.32)	(1.05)	(1.03)	(0.34)	(0.51)	(1.58)
unidirectional*clubmate's outcome	-0.055	-0.039	0.480**	-0.043	0.682	0.438**	-0.280
	(0.34)	(0.06)	(0.23)	(0.23)	(0.49)	(0.21)	(0.23)
own age	-0.385	-0.223	0.521	-0.488	0.540	-0.021	0.584
	(0.42)	(0.22)	(0.37)	(0.31)	(0.51)	(0.42)	(0.36)
own highest high school level	0.851	0.639*	1.554**	-1.099*	1.297	-1.273*	0.508
	(0.74)	(0.36)	(0.67)	(0.63)	(1.44)	(0.72)	(0.69)
own income	0.941	0.057	-1.296***	-0.060	0.935	0.064	1.116*
	(0.67)	(0.29)	(0.47)	(0.47)	(0.79)	(0.58)	(0.66)
Obs	594	624	592	654	650	632	616
R-squared	0.063	0.786	0.218	0.117	0.208	0.336	0.075
Total friend effect=0 (p-value)	0.972	0.111	0.024	0.619	0.606	0.045	0.131
Clustered standard errors	YES	YES	YES	YES	YES	YES	YES
Method	OLOG	OLS	OLOG	OLOG	OLOG	OLOG	OLOG

Note: Estimations for unidirectional friendship from self. Dependent variables are own current happiness (0=very unhappy,...,10=very happy), BMI, sports frequency (1=never,...,6=more than 5 times per week), alcohol quantity (1=never,...,6=10 or more glasses), relationship status (1=currently in a relationship, 0=otherwise) and number of bed partners (1=0 bed partners,...,5=10 or more) and average grade (1-10), respectively. Ordered logistic regression results ('OLOG') are reported, except for BMI, for which linear regression results ('OLS') are reported. Standard errors are clustered at the individual level and are given in parentheses. *** indicates significance at the 0.01 level; ** indicates significance at the 0.05 level; * indicates significance at the 0.10 level. For ordered logit estimations, the pseudo R-squared is reported. 'Total friend effect=0 (p-value)' indicates the p-value of a Wald test with null hypothesis *clubmate's outcome + friend*clubmate's outcome=0.*

Table 4.C3: OLS or ordered logistic regression of own current happiness and behaviour on own lagged and clubmate's current and lagged happiness and behaviour (mutual friendship)

	happiness	bmi	sports	alcohol	relation	bed partners	grade
constant		4.808					
		(4.19)					
own outcome (lag)	0.579	0.932***	1.530***	1.008***	2.257***	3.160***	0.509
	(0.49)	(0.07)	(0.38)	(0.27)	(0.81)	(0.69)	(0.47)
clubmate's outcome (lag)	-0.195	-0.140**	0.124**	0.032	0.185	0.117	0.225*
	(0.12)	(0.05)	(0.05)	(0.07)	(0.25)	(0.12)	(0.12)
clubmate's outcome	-0.091	0.160**	0.116	-0.145	-0.485***	0.028	0.002
	(0.11)	(0.06)	(0.13)	(0.10)	(0.10)	(0.14)	(0.09)
mutual	-6.490***	0.641	-0.019	-2.072**	-0.184	0.073	3.649**
	(2.08)	(1.23)	(0.97)	(0.97)	(0.22)	(0.30)	(1.50)
mutual*clubmate's outcome	0.820***	-0.030	-0.025	0.424**	0.219	0.061	-0.522**
	(0.26)	(0.06)	(0.24)	(0.20)	(0.59)	(0.15)	(0.22)
own age	-0.360	-0.221	0.534	-0.486	0.531	-0.031	0.582
	(0.43)	(0.22)	(0.37)	(0.32)	(0.52)	(0.42)	(0.36)
own highest high school level	0.828	0.633*	1.551**	-1.096*	1.249	-1.275*	0.491
	(0.72)	(0.36)	(0.68)	(0.63)	(1.43)	(0.73)	(0.70)
own income	0.935	0.051	-1.293***	-0.049	0.896	0.049	1.136*
	(0.66)	(0.29)	(0.47)	(0.47)	(0.79)	(0.58)	(0.66)
Obs	594	624	592	654	650	632	616
R-squared	0.068	0.785	0.216	0.120	0.197	0.334	0.077
Total friend effect=0 (p-value)	0.001	0.043	0.664	0.076	0.614	0.554	0.012
Clustered standard errors	YES	YES	YES	YES	YES	YES	YES
Method	OLOG	OLS	OLOG	OLOG	OLOG	OLOG	OLOG

Note: Estimations for mutual friendship. Dependent variables are own current happiness (0=very unhappy,...,10=very happy), BMI, sports frequency (1=never,...,6=more than 5 times per week), alcohol quantity (1=never,...,6=10 or more glasses), relationship status (1=currently in a relationship, 0=otherwise) and number of bed partners (1=0 bed partners,...,5=10 or more) and average grade (1-10), respectively. Ordered logistic regression results ('OLOG') are reported, except for BMI, for which linear regression results ('OLS') are reported. Standard errors are clustered at the individual level and are given in parentheses. *** indicates significance at the 0.01 level; ** indicates significance at the 0.05 level; * indicates significance at the 0.10 level. For ordered logit estimations, the pseudo R-squared is reported. 'Total friend effect=0 (p-value)' indicates the p-value of a Wald test with null hypothesis *clubmate's outcome + friend*clubmate's outcome=0.*

Table 4.C4: OLS or ordered logistic regression of own current happiness and behaviour on own lagged and clubmate's current and lagged happiness and behaviour (missing observations imputed)

	happiness	bmi	sports	alcohol	relation	bed partners	grade
constant		4.645					
		(4.15)					
own outcome (lag)	0.562	0.933***	1.527***	1.004***	2.298***	3.162***	0.503
	(0.49)	(0.07)	(0.38)	(0.27)	(0.81)	(0.70)	(0.47)
clubmate's outcome (lag)	-0.197	-0.139**	0.126**	0.039	0.198	0.097	0.221*
	(0.12)	(0.05)	(0.05)	(0.07)	(0.25)	(0.11)	(0.12)
clubmate's outcome	-0.102	0.163**	0.041	-0.152	-0.499***	-0.027	0.057
	(0.12)	(0.06)	(0.13)	(0.11)	(0.12)	(0.14)	(0.10)
friend	-4.442***	0.898	-1.032*	-1.548**	0.256	-0.387	3.430***
	(1.59)	(0.86)	(0.60)	(0.76)	(0.21)	(0.32)	(1.24)
friend*clubmate's outcome	0.561***	-0.037	0.256*	0.288*	0.272	0.258*	-0.500***
	(0.20)	(0.04)	(0.14)	(0.16)	(0.43)	(0.14)	(0.19)
own age	-0.357	-0.218	0.526	-0.478	0.544	-0.030	0.577
	(0.43)	(0.22)	(0.37)	(0.32)	(0.52)	(0.42)	(0.36)
own highest high school level	0.802	0.626*	1.559**	-1.100*	1.247	-1.265*	0.506
	(0.72)	(0.36)	(0.68)	(0.64)	(1.43)	(0.71)	(0.70)
own income	0.941	0.047	-1.298***	-0.045	0.892	0.063	1.130*
	(0.67)	(0.29)	(0.47)	(0.47)	(0.78)	(0.58)	(0.66)
Obs	594	624	592	654	650	632	616
R-squared	0.066	0.786	0.217	0.119	0.200	0.335	0.078
Total friend effect=0 (p-value)	0.006	0.023	0.025	0.239	0.511	0.108	0.005
Clustered standard errors	YES	YES	YES	YES	YES	YES	YES
Method	OLOG	OLS	OLOG	OLOG	OLOG	OLOG	OLOG

Note: Estimations with missing observations imputed. Dependent variables are own current happiness (0=very unhappy,...,10=very happy), BMI, sports frequency (1=never,...,6=more than 5 times per week), alcohol quantity (1=never,...,6=10 or more glasses), relationship status (1=currently in a relationship, 0=otherwise) and number of bed partners (1=0 bed partners,...,5=10 or more) and average grade (1-10), respectively. Ordered logistic regression results ('OLOG') are reported, except for BMI, for which linear regression results ('OLS') are reported. Standard errors are clustered at the individual level and are given in parentheses. *** indicates significance at the 0.01 level; ** indicates significance at the 0.05 level; * indicates significance at the 0.10 level. For ordered logit estimations, the pseudo R-squared is reported rather than the R-squared. 'Total friend effect=0 (p-value)' indicates the p-value of a Wald test with null hypothesis *clubmate's outcome + friend*clubmate's outcome=0*. Missing values of the outcome variables are imputed. A dummy variable ('imputed observation') is included, which equals a 1 if a value was imputed, 0 otherwise. There are zero missing values for sports frequency, alcohol quantity, bed partners, and relationship status; the dummy is therefore omitted for these outcome variables.

Table 4.C5: OLS or ordered logistic regression of own current happiness and behaviour on own lagged and clubmate's current and lagged happiness and behaviour (including outliers)

	happiness	bmi	sports	alcohol	relation	bed partners	grade
constant		4.645					
		(4.15)					
own outcome (lag)	0.550	0.933***	1.514***	0.987***	2.298***	3.162***	0.503
	(0.50)	(0.07)	(0.38)	(0.27)	(0.81)	(0.70)	(0.47)
clubmate's outcome (lag)	-0.171	-0.139**	0.077*	0.061	0.198	0.097	0.221*
	(0.11)	(0.05)	(0.05)	(0.07)	(0.25)	(0.11)	(0.12)
clubmate's outcome	0.010	0.163**	-0.055	-0.176	-0.499***	-0.027	0.057
	(0.10)	(0.06)	(0.12)	(0.11)	(0.12)	(0.14)	(0.10)
friend	-1.622	0.898	-0.952*	-1.471*	0.256	-0.387	3.430***
	(1.65)	(0.86)	(0.50)	(0.77)	(0.21)	(0.32)	(1.24)
friend*clubmate's outcome	0.211	-0.037	0.251**	0.272*	0.272	0.258*	-0.500***
	(0.21)	(0.04)	(0.12)	(0.16)	(0.43)	(0.14)	(0.19)
own age	-0.375	-0.218	0.551	-0.460	0.544	-0.030	0.577
	(0.43)	(0.22)	(0.37)	(0.30)	(0.52)	(0.42)	(0.36)
own highest high school level	0.868	0.626*	1.525**	-1.136*	1.247	-1.265*	0.506
	(0.71)	(0.36)	(0.67)	(0.63)	(1.43)	(0.71)	(0.70)
own income	0.968	0.047	-1.318***	-0.054	0.892	0.063	1.130*
	(0.67)	(0.29)	(0.48)	(0.47)	(0.78)	(0.58)	(0.66)
Obs	632	624	630	658	650	632	616
R-squared	0.065	0.786	0.217	0.118	0.200	0.335	0.078
Total friend effect=0 (p-value)	0.215	0.023	0.070	0.393	0.511	0.108	0.005
Clustered standard errors	YES	YES	YES	YES	YES	YES	YES
Method	OLOG	OLS	OLOG	OLOG	OLOG	OLOG	OLOG

Note: Estimations including outliers. Dependent variables are own current happiness (0=very unhappy,...,10=very happy), BMI, sports frequency (1=never,...,6=more than 5 times per week), alcohol quantity (1=never,...,6=10 or more glasses), relationship status (1=currently in a relationship, 0=otherwise) and number of bed partners (1=0 bed partners,...,5=10 or more) and average grade (1-10), respectively. Ordered logistic regression results ('OLOG') are reported, except for BMI, for which linear regression results ('OLS') are reported. Standard errors are clustered at the individual level and are given in parentheses. *** indicates significance at the 0.01 level; ** indicates significance at the 0.05 level; * indicates significance at the 0.10 level. For ordered logit estimations, the pseudo R-squared is reported. 'Total friend effect=0 (p-value)' indicates the p-value of a Wald test with null hypothesis *clubmate's outcome + friend*clubmate's outcome=0*.

Table 4.C6: OLS or ordered logistic regression of own current happiness and behaviour on own lagged and clubmate's current and lagged happiness and behaviour (excluding individuals who knew each other before)

	happiness	bmi	sports	alcohol	relation	bed partners	grade
constant		4.212					
		(4.08)					
own outcome (lag)	0.536	0.932***	1.527***	1.001***	2.417***	3.129***	0.524
	(0.47)	(0.07)	(0.38)	(0.28)	(0.82)	(0.70)	(0.47)
clubmate's outcome (lag)	-0.173	-0.143***	0.162***	0.024	0.072	0.098	0.266**
	(0.12)	(0.05)	(0.06)	(0.07)	(0.24)	(0.11)	(0.12)
clubmate's outcome	-0.074	0.167**	0.022	-0.132	-0.474***	-0.021	0.037
	(0.12)	(0.06)	(0.13)	(0.11)	(0.16)	(0.14)	(0.10)
friend	-4.311***	0.610	-0.685	-1.568**	0.194	-0.328	4.272***
	(1.62)	(0.93)	(0.63)	(0.78)	(0.25)	(0.33)	(1.31)
friend*clubmate's outcome	0.537***	-0.025	0.173	0.302*	0.569	0.242*	-0.623***
	(0.20)	(0.04)	(0.14)	(0.17)	(0.45)	(0.14)	(0.20)
own age	-0.298	-0.195	0.556	-0.506	0.652	-0.045	0.596*
	(0.42)	(0.22)	(0.38)	(0.32)	(0.55)	(0.42)	(0.36)
own highest high school level	0.821	0.614*	1.483**	-1.108*	1.242	-1.238*	0.584
	(0.73)	(0.36)	(0.67)	(0.65)	(1.43)	(0.72)	(0.70)
own income	0.892	0.047	-1.318***	-0.015	0.802	0.096	1.161*
	(0.66)	(0.30)	(0.47)	(0.47)	(0.80)	(0.59)	(0.65)
Obs	575	602	570	632	628	610	594
R-squared	0.059	0.786	0.215	0.121	0.214	0.331	0.085
Total friend effect=0 (p-value)	0.008	0.016	0.133	0.161	0.784	0.129	0.001
Clustered standard errors	YES	YES	YES	YES	YES	YES	YES
Method	OLOG	OLS	OLOG	OLOG	OLOG	OLOG	OLOG

Note: Estimations excluding pairs of previous acquaintances. Dependent variables are own current happiness (0=very unhappy,...,10=very happy), BMI, sports frequency (1=never,...,6=more than 5 times per week), alcohol quantity (1=never,...,6=10 or more glasses), relationship status (1=currently in a relationship, 0=otherwise) and number of bed partners (1=0 bed partners,...,5=10 or more) and average grade (1-10), respectively. Ordered logistic regression results ('OLOG') are reported, except for BMI, for which linear regression results ('OLS') are reported. Standard errors are clustered at the individual level and are given in parentheses. *** indicates significance at the 0.01 level; ** indicates significance at the 0.05 level; * indicates significance at the 0.10 level. For ordered logit estimations, the pseudo R-squared is reported. 'Total friend effect=0 (p-value)' indicates the p-value of a Wald test with null hypothesis *clubmate's outcome + friend*clubmate's outcome=0*.

4.C.2 By outcome variable: different sets of regressors

Table 4.C7: OLS and ordered logistic regression of own current happiness on own lagged happiness and clubmate's current and lagged happiness

	ols0	ols1	ols2	ols3	ols4	ologit0	ologit1	ologit2	ologit3	ologit4
	b/se	b/se	b/se	b/se	b/se	b/se	b/se	b/se	b/se	b/se
constant	7.802***	4.862***	4.862**	5.457**	8.763**					
	(0.47)	(0.70)	(2.08)	(2.11)	(3.60)					
clubmate's happiness	-0.006	0.017	0.017	-0.058	-0.034	-0.002	0.038	0.038	-0.140	-0.102
	(0.06)	(0.05)	(0.04)	(0.05)	(0.05)	(0.10)	(0.11)	(0.09)	(0.11)	(0.12)
own lag happiness		0.443***	0.443	0.446*	0.334		0.677***	0.677	0.697	0.562
		(0.05)	(0.27)	(0.26)	(0.21)		(0.11)	(0.52)	(0.52)	(0.49)
clubmate's lag happiness		-0.084	-0.084	-0.086	-0.102*		-0.146	-0.146	-0.158	-0.197
		(0.05)	(0.05)	(0.05)	(0.05)		(0.10)	(0.11)	(0.11)	(0.12)
friend				-2.408**	-1.577*				-5.536***	-4.442***
				(0.92)	(0.83)				(1.69)	(1.59)
friend*clubmate's happiness				0.305***	0.204**				0.694***	0.561***
				(0.11)	(0.10)				(0.21)	(0.20)
own age					-0.247					-0.357
					(0.20)					(0.43)
own highest high school level					0.417					0.802
					(0.36)					(0.72)
own hh relative income					0.459					0.941
					(0.30)					(0.67)
Obs	594	594	594	594	594	594	594	594	594	594
R-squared	0.000	0.125	0.125	0.134	0.227	0.000	0.028	0.028	0.033	0.066
Total friend effect=0 (p-value)				0.012	0.054				0.002	0.006
Clustered standard errors	NO	NO	YES	YES	YES	NO	NO	YES	YES	YES
Friend	NO	NO	NO	YES	YES	NO	NO	NO	YES	YES
Controls	NO	NO	NO	NO	YES	NO	NO	NO	NO	YES

Note: Dependent variable is own current happiness (0=very unhappy,...,10=very happy). Standard errors are given in parentheses. In case of clustering, standard errors are clustered at the individual level. *** indicates significance at the 0.01 level; ** indicates significance at the 0.05 level; * indicates significance at the 0.10 level. For ordered logit estimations, the pseudo R-squared is reported rather than the R-squared. 'Total friend effect=0 (p-value)' indicates the p-value of a Wald test with null hypothesis *clubmate's outcome + friend*clubmate's outcome=0.*

Table 4.C8: OLS and ordered logistic regression of own current BMI on own lagged BMI and clubmate's current and lagged BMI

	ols0	ols1	ols2	ols3	ols4
constant	20.804***	1.414**	1.414	1.038	4.645
	(0.87)	(0.62)	(1.29)	(1.42)	(4.15)
clubmate's BMI	0.032	0.154***	0.154**	0.167**	0.163**
	(0.04)	(0.04)	(0.06)	(0.07)	(0.06)
own lag BMI		0.929***	0.929***	0.931***	0.933***
		(0.02)	(0.06)	(0.06)	(0.07)
clubmate's lag BMI		-0.138***	-0.138**	-0.138**	-0.139**
		(0.04)	(0.06)	(0.06)	(0.05)
friend				1.341	0.898
				(1.02)	(0.86)
friend*clubmate's BMI				-0.057	-0.037
				(0.05)	(0.04)
own age					-0.218
					(0.22)
own highest high school level					0.626*
					(0.36)
own hh relative income					0.047
					(0.29)
Obs	624	624	624	624	624
R-squared	0.001	0.775	0.775	0.776	0.786
Total friend effect=0 (p-value)				0.057	0.023
Clustered standard errors	NO	NO	YES	YES	YES
Friend	NO	NO	NO	YES	YES
Controls	NO	NO	NO	NO	YES

Note: Dependent variable is own current BMI. Standard errors are given in parentheses. In case of clustering, standard errors are clustered at the individual level. *** indicates significance at the 0.01 level; ** indicates significance at the 0.05 level; * indicates significance at the 0.10 level. For ordered logit estimations, the pseudo R-squared is reported rather than the R-squared. 'Total friend effect=0 (p-value)' indicates the p-value of a Wald test with null hypothesis *clubmate's outcome + friend*clubmate's outcome=0.*

Table 4.C9: OLS and ordered logistic regression of own current sports frequency on own lagged sports frequency and clubmate's current and lagged sports frequency

	ols0	ols1	ols2	ols3	ols4	ologit0	ologit1	ologit2	ologit3	ologit4
	b/se	b/se	b/se	b/se	b/se	b/se	b/se	b/se	b/se	b/se
constant	3.606***	1.434***	1.434**	1.541**	-0.990					
	(0.20)	(0.24)	(0.61)	(0.62)	(2.74)					
clubmate's sports frequency	0.057	0.031	0.031	0.003	0.008	0.116	0.106	0.106	0.044	0.041
	(0.05)	(0.05)	(0.05)	(0.06)	(0.05)	(0.09)	(0.10)	(0.11)	(0.11)	(0.13)
own lag sports frequency		0.585***	0.585***	0.584***	0.642***		1.154***	1.154***	1.154***	1.527***
		(0.04)	(0.14)	(0.14)	(0.14)		(0.09)	(0.30)	(0.30)	(0.38)
clubmate's lag sports frequency		0.035	0.035	0.036	0.046**		0.074	0.074	0.077*	0.126**
		(0.04)	(0.02)	(0.02)	(0.02)		(0.09)	(0.05)	(0.05)	(0.05)
friend				-0.405	-0.393				-0.910	-1.032*
				(0.31)	(0.24)				(0.69)	(0.60)
friend*clubmate's sports frequency				0.101	0.098*				0.228	0.256*
				(0.08)	(0.06)				(0.18)	(0.14)
own age					0.189					0.526
					(0.13)					(0.37)
own highest high school level					0.713**					1.559**
					(0.28)					(0.68)
own hh relative income					-0.473***					-1.298***
					(0.17)					(0.47)
Obs	620	592	592	592	592	620	592	592	592	592
R-squared	0.002	0.300	0.300	0.301	0.478	0.001	0.107	0.107	0.108	0.217
Total friend effect=0 (p-value)				0.146	0.048				0.116	0.025
Clustered standard errors	NO	NO	YES	YES	YES	NO	NO	YES	YES	YES
Friend	NO	NO	NO	YES	YES	NO	NO	NO	YES	YES
Controls	NO	NO	NO	NO	YES	NO	NO	NO	NO	YES

Note: Dependent variable is own current frequency of doing sports (1=never, 2=less than once a month, 3=1-3 times a month, 4= 1-2 times per week, 5=3-4 times per week, and 6=more than 5 times per week). Standard errors are given in parentheses. In case of clustering, standard errors are clustered at the individual level. *** indicates significance at the 0.01 level; ** indicates significance at the 0.05 level; * indicates significance at the 0.10 level. For ordered logit estimations, the pseudo R-squared is reported rather than the R-squared. 'Total friend effect=0 (p-value)' indicates the p-value of a Wald test with null hypothesis *clubmate's outcome + friend*clubmate's outcome=0.*

Table 4.C10: OLS and ordered logistic regression of own current alcohol quantity on own lagged alcohol quantity and clubmate's current and lagged alcohol quantity

	ols0	ols1	ols2	ols3	ols4	ologit0	ologit1	ologit2	ologit3	ologit4
	b/se	b/se	b/se	b/se	b/se	b/se	b/se	b/se	b/se	b/se
constant	4.760***	2.782***	2.782***	3.019***	8.113**					
	(0.19)	(0.25)	(0.52)	(0.55)	(3.14)					
clubmate's alcohol quantity	-0.022	-0.015	-0.015	-0.058	-0.068	-0.034	-0.025	-0.025	-0.110	-0.152
	(0.04)	(0.04)	(0.04)	(0.05)	(0.05)	(0.07)	(0.08)	(0.08)	(0.10)	(0.11)
own lag alcohol quantity		0.461***	0.461***	0.457***	0.475***		0.936***	0.936***	0.933***	1.004***
		(0.04)	(0.12)	(0.11)	(0.11)		(0.08)	(0.28)	(0.28)	(0.27)
clubmate's lag alcohol quantity		0.014	0.014	0.017	0.034		0.006	0.006	0.013	0.039
		(0.04)	(0.04)	(0.04)	(0.04)		(0.08)	(0.07)	(0.07)	(0.07)
friend				-0.892**	-0.784**				-1.668**	-1.548**
				(0.38)	(0.38)				(0.67)	(0.76)
friend*clubmate's alcohol quantity				0.158**	0.137*				0.305**	0.288*
				(0.08)	(0.07)				(0.14)	(0.16)
own age					-0.230					-0.478
					(0.14)					(0.32)
own highest high school level					-0.555*					-1.100*
					(0.29)					(0.64)
own hh relative income					-0.063					-0.045
					(0.22)					(0.47)
Obs	654	654	654	654	654	654	654	654	654	654
R-squared	0.000	0.187	0.187	0.196	0.279	0.000	0.079	0.079	0.083	0.119
Total friend effect=0 (p-value)				0.054	0.187				0.052	0.239
Clustered standard errors	NO	NO	YES	YES	YES	NO	NO	YES	YES	YES
Friend	NO	NO	NO	YES	YES	NO	NO	NO	YES	YES
Controls	NO	NO	NO	NO	YES	NO	NO	NO	NO	YES

Note: Dependent variable is own current quantity alcohol consumed on a drinking night (1=never, 2=1-2 glasses, 3=3-4 glasses, 4= 5-6 glasses, 5=7-9 glasses, and 6=10 or more glasses). Standard errors are given in parentheses. In case of clustering, standard errors are clustered at the individual level. *** indicates significance at the 0.01 level; ** indicates significance at the 0.05 level; * indicates significance at the 0.10 level. For ordered logit estimations, the pseudo R-squared is reported rather than the R-squared. 'Total friend effect=0 (p-value)' indicates the p-value of a Wald test with null hypothesis *clubmate's outcome + friend*clubmate's outcome=0*.

Table 4.C11: OLS and ordered logistic regression of own current relationship status on own lagged relationship status and clubmate's current and lagged relationship status

	ols0	ols1	ols2	ols3	ols4	ologit0	ologit1	ologit2	ologit3	ologit4
	b/se	b/se	b/se	b/se	b/se	b/se	b/se	b/se	b/se	b/se
constant	0.306***	0.180***	0.180***	0.167***	-2.039					
	(0.02)	(0.02)	(0.06)	(0.06)	(1.61)					
clubmate's in serious relationship	-0.055	-0.079**	-0.079***	-0.086***	-0.076***	-0.274	-0.451**	-0.451***	-0.504***	-0.499***
	(0.04)	(0.04)	(0.02)	(0.02)	(0.02)	(0.19)	(0.23)	(0.06)	(0.12)	(0.12)
own lag in serious relationship		0.390***	0.390***	0.394***	0.422***		1.850***	1.850***	1.878***	2.298***
		(0.04)	(0.14)	(0.14)	(0.13)		(0.19)	(0.68)	(0.68)	(0.81)
clubmate's lag in serious relationship		0.033	0.033	0.034	0.039		0.172	0.172	0.179	0.198
		(0.04)	(0.04)	(0.04)	(0.04)		(0.22)	(0.24)	(0.24)	(0.25)
friend				0.040	0.040				0.233	0.256
				(0.04)	(0.04)				(0.21)	(0.21)
friend*clubmate's in serious relationship				0.047	0.030				0.287	0.272
				(0.08)	(0.07)				(0.47)	(0.43)
own age					0.081					0.544
					(0.09)					(0.52)
own highest high school level					0.151					1.247
					(0.15)					(1.43)
own hh relative income					0.117					0.892
					(0.11)					(0.78)
Obs	658	650	650	650	650	658	650	650	650	650
R-squared	0.003	0.165	0.165	0.168	0.223	0.003	0.133	0.133	0.136	0.200
Total friend effect=0 (p-value)				0.551	0.452				0.557	0.511
Clustered standard errors	NO	NO	YES	YES	YES	NO	NO	YES	YES	YES
Friend	NO	NO	NO	YES	YES	NO	NO	NO	YES	YES
Controls	NO	NO	NO	NO	YES	NO	NO	NO	NO	YES

Note: Dependent variable is own current relationship status (1=currently in a relationship, 0=otherwise). Standard errors are given in parentheses. In case of clustering, standard errors are clustered at the individual level. *** indicates significance at the 0.01 level; ** indicates significance at the 0.05 level; * indicates significance at the 0.10 level. For ordered logit estimations, the pseudo R-squared is reported rather than the R-squared. 'Total friend effect=0 (p-value)' indicates the p-value of a Wald test with null hypothesis *clubmate's outcome + friend*clubmate's outcome=0.*

Table 4.C12: OLS and ordered logistic regression of own current number of bed partners on own lagged bed partners and clubmate's current and lagged bed partners

	ols0	ols1	ols2	ols3	ols4	ologit0	ologit1	ologit2	ologit3	ologit4
	b/se	b/se	b/se	b/se	b/se	b/se	b/se	b/se	b/se	b/se
constant	1.765***	0.732***	0.732***	0.790***	0.785					
	(0.09)	(0.07)	(0.15)	(0.16)	(2.93)					
clubmate's bed partners	0.097**	0.047	0.047	0.023	0.005	0.132**	0.071	0.071	0.008	-0.027
	(0.04)	(0.04)	(0.05)	(0.05)	(0.05)	(0.06)	(0.11)	(0.14)	(0.15)	(0.14)
own lag bed partners		0.950***	0.950***	0.945***	0.962***		3.088***	3.088***	3.087***	3.162***
		(0.03)	(0.08)	(0.08)	(0.08)		(0.19)	(0.68)	(0.68)	(0.70)
clubmate's lag bed partners		0.010	0.010	0.001	0.026		0.070	0.070	0.047	0.097
		(0.05)	(0.05)	(0.05)	(0.04)		(0.13)	(0.13)	(0.12)	(0.11)
friend				-0.189	-0.201				-0.367	-0.387
				(0.14)	(0.14)				(0.33)	(0.32)
friend*clubmate's bed partners				0.102	0.102*				0.250	0.258*
				(0.06)	(0.06)				(0.15)	(0.14)
own age					-0.007					-0.030
					(0.13)					(0.42)
own highest high school level					-0.534*					-1.265*
					(0.28)					(0.71)
own hh relative income					0.135					0.063
					(0.18)					(0.58)
Obs	658	632	632	632	632	658	632	632	632	632
R-squared	0.009	0.584	0.584	0.587	0.622	0.003	0.313	0.313	0.315	0.335
Total friend effect=0 (p-value)				0.064	0.058				0.116	0.108
Clustered standard errors	NO	NO	YES	YES	YES	NO	NO	YES	YES	YES
Friend	NO	NO	NO	YES	YES	NO	NO	NO	YES	YES
Controls	NO	NO	NO	NO	YES	NO	NO	NO	NO	YES

Note: Dependent variable is own current number of bed partners (1=0 bed partners, 2=1-3, 3=4-6, 4=7-9, 5=10 or more). Standard errors are given in parentheses. In case of clustering, standard errors are clustered at the individual level. *** indicates significance at the 0.01 level; ** indicates significance at the 0.05 level; * indicates significance at the 0.10 level. For ordered logit estimations, the pseudo R-squared is reported rather than the R-squared. 'Total friend effect=0 (p-value)' indicates the p-value of a Wald test with null hypothesis *clubmate's outcome + friend*clubmate's outcome=0*.

Table 4.C13: OLS and ordered logistic regression of own current grade on own lagged grade and clubmate's current and lagged grade

	ols0	ols1	ols2	ols3	ols4	ologit0	ologit1	ologit2	ologit3	ologit4
	b/se	b/se	b/se	b/se	b/se	b/se	b/se	b/se	b/se	b/se
constant	6.749***	4.833***	4.833***	4.513***	-1.316					
	(0.27)	(0.48)	(1.28)	(1.29)	(3.20)					
clubmate's grade	-0.008	-0.036	-0.036	0.014	0.024	-0.047	-0.112	-0.112	0.020	0.057
	(0.04)	(0.04)	(0.03)	(0.04)	(0.03)	(0.10)	(0.11)	(0.08)	(0.10)	(0.10)
own lag grade		0.213***	0.213	0.209	0.184		0.503***	0.503	0.495	0.503
		(0.05)	(0.16)	(0.16)	(0.16)		(0.12)	(0.42)	(0.42)	(0.47)
clubmate's lag grade		0.085*	0.085*	0.084*	0.084**		0.183	0.183	0.188	0.221*
		(0.05)	(0.05)	(0.05)	(0.04)		(0.12)	(0.12)	(0.12)	(0.12)
friend				1.332**	1.185**				3.410***	3.430***
				(0.54)	(0.47)				(1.23)	(1.24)
friend*clubmate's grade				-0.188**	-0.171**				-0.485***	-0.500***
				(0.08)	(0.07)				(0.19)	(0.19)
own age					0.214*					0.577
					(0.11)					(0.36)
own highest high school level					0.140					0.506
					(0.23)					(0.70)
own hh relative income					0.393**					1.130*
					(0.19)					(0.66)
Obs	616	616	616	616	616	616	616	616	616	616
R-squared	0.000	0.041	0.041	0.050	0.173	0.000	0.015	0.015	0.019	0.078
Total friend effect=0 (p-value)				0.007	0.011				0.002	0.005
Clustered standard errors	NO	NO	YES	YES	YES	NO	NO	YES	YES	YES
Friend	NO	NO	NO	YES	YES	NO	NO	NO	YES	YES
Controls	NO	NO	NO	NO	YES	NO	NO	NO	NO	YES

Note: Dependent variable is own current average grade (1-10). Standard errors are given in parentheses. In case of clustering, standard errors are clustered at the individual level. *** indicates significance at the 0.01 level; ** indicates significance at the 0.05 level; * indicates significance at the 0.10 level. For ordered logit estimations, the pseudo R-squared is reported rather than the R-squared. 'Total friend effect=0 (p-value)' indicates the p-value of a Wald test with null hypothesis *clubmate's outcome + friend*clubmate's outcome=0*.

4.C.3 By outcome variable: non-linear specification

Table 4.C14: Ordered logistic regression of own current happiness on categories of own lagged happiness and clubmate's current and lagged happiness

	ologit1	ologit2	ologit3	ologit4	ologit5
5.clubmate's happiness	-0.464	-0.217	-0.217	0.947	0.861
	(0.72)	(0.73)	(0.70)	(0.73)	(0.68)
6.clubmate's happiness	-0.476	-0.421	-0.421	0.267	0.133
	(0.72)	(0.73)	(0.71)	(0.80)	(0.76)
7.clubmate's happiness	-0.346	0.007	0.007	0.913	0.830
	(0.51)	(0.55)	(0.51)	(0.59)	(0.56)
8.clubmate's happiness	-0.338	0.036	0.036	0.765	0.756
	(0.51)	(0.54)	(0.53)	(0.64)	(0.62)
9.clubmate's happiness	-0.352	0.085	0.085	0.632	0.625
	(0.52)	(0.56)	(0.53)	(0.64)	(0.61)
own lag happiness		0.686***	0.686	0.705	0.567
		(0.11)	(0.52)	(0.51)	(0.49)
clubmate's lag happiness		-0.148	-0.148	-0.164	-0.203*
		(0.11)	(0.11)	(0.11)	(0.12)
friend				1.651*	1.436
				(0.90)	(1.07)
5.clubmate's happiness*friend				-4.238***	-4.139***
				(1.11)	(1.25)
7.clubmate's happiness*friend				-2.291***	-1.837*
				(0.86)	(1.03)
8.clubmate's happiness*friend				-1.690*	-1.519
				(0.94)	(1.11)
9.clubmate's happiness*friend				-0.851	-0.647
				(0.97)	(1.15)
own age					-0.383
					(0.44)
own highest high school level					0.819
					(0.71)
own hh relative income					0.959
					(0.67)
Obs	632	632	632	632	632
Pseudo R-squared	0.000	0.029	0.029	0.039	0.072
Clustered standard errors	NO	NO	YES	YES	YES
Friend	NO	NO	NO	YES	YES
Controls	NO	NO	NO	NO	YES

Note: Dependent variable is own current happiness (0=very unhappy,...,10=very happy). Standard errors are given in parentheses. In case of clustering, standard errors are clustered at the individual level. *** indicates significance at the 0.01 level; ** indicates significance at the 0.05 level; * indicates significance at the 0.10 level. The first digit of the coefficient name indicates the category number. Outliers are included.

Table 4.C15: Ordered logistic regression of own current sports frequency on categories of own lagged sports frequency and clubmate's current and lagged sports frequency

	ologit1	ologit2	ologit3	ologit4	ologit5
2.clubmate's sports frequency	-0.162	-0.189	-0.189	-0.249	-0.528
	(0.46)	(0.49)	(0.38)	(0.49)	(0.49)
3.clubmate's sports frequency	-0.298	-0.267	-0.267	-0.476	-0.760*
	(0.41)	(0.46)	(0.32)	(0.42)	(0.42)
4.clubmate's sports frequency	-0.020	0.024	0.024	-0.166	-0.739
	(0.39)	(0.45)	(0.46)	(0.57)	(0.53)
5.clubmate's sports frequency	0.045	0.026	0.026	-0.280	-0.513
	(0.40)	(0.48)	(0.43)	(0.51)	(0.49)
6.clubmate's sports frequency	-0.149	-0.214	-0.214	-0.833	-1.872**
	(0.63)	(0.71)	(0.55)	(0.82)	(0.74)
own lag sports frequency		1.143***	1.143***	1.142***	1.528***
		(0.09)	(0.30)	(0.30)	(0.37)
clubmate's lag sports frequency		0.057	0.057	0.056	0.127**
		(0.09)	(0.05)	(0.05)	(0.06)
friend				-0.605	-0.481
				(0.82)	(0.81)
2.clubmate's sports frequency*friend				-0.051	0.057
				(1.06)	(1.10)
3.clubmate's sports frequency*friend				0.604	0.000
				(0.85)	(0.90)
4.clubmate's sports frequency*friend				0.532	0.614
				(0.87)	(0.89)
5.clubmate's sports frequency*friend				0.906	0.602
				(0.89)	(0.85)
6.clubmate's sports frequency*friend				1.458	2.166**
				(1.15)	(1.06)
own age					0.566
					(0.37)
own highest high school level					1.592**
					(0.68)
own hh relative income					-1.337***
					(0.47)
Obs	658	630	630	630	630
Pseudo R-squared	0.001	0.108	0.108	0.109	0.221
Clustered standard errors	NO	NO	YES	YES	YES
Friend	NO	NO	NO	YES	YES
Controls	NO	NO	NO	NO	YES

Note: Dependent variable is own current frequency of doing sports (1=never, 2=less than once a month, 3=1-3 times a month, 4= 1-2 times per week, 5=3-4 times per week, and 6=more than 5 times per week). Standard errors are given in parentheses. In case of clustering, standard errors are clustered at the individual level. *** indicates significance at the 0.01 level; ** indicates significance at the 0.05 level; * indicates significance at the 0.10 level. The first digit of the coefficient name indicates the category number. Outliers are included.

Table 4.C16: Ordered logistic regression of own current alcohol quantity on categories of own lagged alcohol quantity and clubmate's current and lagged alcohol quantity

	ologit1	ologit2	ologit3	ologit4	ologit5
3.clubmate's alcohol quantity	0.486	-0.435	-0.435	-1.888***	-2.035***
	(0.81)	(0.92)	(0.94)	(0.71)	(0.70)
4.clubmate's alcohol quantity	0.124	-0.817	-0.817	-2.550***	-2.661***
	(0.79)	(0.91)	(0.93)	(0.73)	(0.74)
5.clubmate's alcohol quantity	0.275	-0.613	-0.613	-2.135***	-2.279***
	(0.80)	(0.91)	(0.92)	(0.72)	(0.71)
6.clubmate's alcohol quantity	0.218	-0.705	-0.705	-2.557***	-2.826***
	(0.79)	(0.90)	(0.93)	(0.71)	(0.75)
own lag alcohol quantity		0.923***	0.923***	0.930***	1.005***
		(0.08)	(0.28)	(0.28)	(0.27)
clubmate's lag alcohol quantity		0.011	0.011	0.015	0.054
		(0.08)	(0.07)	(0.07)	(0.07)
friend				-2.381**	-2.327**
				(0.94)	(1.13)
3.clubmate's alcohol quantity*friend				1.344	1.231
				(1.02)	(1.11)
4.clubmate's alcohol quantity*friend				2.336**	2.387**
				(0.96)	(1.11)
5.clubmate's alcohol quantity*friend				1.651*	1.651
				(0.92)	(1.15)
6.clubmate's alcohol quantity*friend				2.760***	2.682**
				(0.97)	(1.18)
own age					-0.466
					(0.30)
own highest high school level					-1.139*
					(0.64)
own hh relative income					-0.062
					(0.47)
Obs	658	658	658	658	658
Pseudo R-squared	0.002	0.080	0.080	0.087	0.124
Clustered standard errors	NO	NO	YES	YES	YES
Friend	NO	NO	NO	YES	YES
Controls	NO	NO	NO	NO	YES

Note: Dependent variable is own current quantity alcohol consumed on a drinking night (1=never, 2=1-2 glasses, 3=3-4 glasses, 4= 5-6 glasses, 5=7-9 glasses, and 6=10 or more glasses). Standard errors are given in parentheses. In case of clustering, standard errors are clustered at the individual level. *** indicates significance at the 0.01 level; ** indicates significance at the 0.05 level; * indicates significance at the 0.10 level. The first digit of the coefficient name indicates the category number. Outliers are included.

Table 4.C17: Ordered logistic regression of own current relationship status on categories of own lagged relationship status and clubmate's current and lagged relationship status

	ologit1	ologit2	ologit3	ologit4	ologit5
1.clubmate's now in serious relationship	-0.274	-0.451**	-0.451***	-0.504***	-0.499***
	(0.19)	(0.23)	(0.06)	(0.12)	(0.12)
own lag now in serious relationship		1.850***	1.850***	1.878***	2.298***
		(0.19)	(0.68)	(0.68)	(0.81)
clubmate's lag now in serious relationship		0.172	0.172	0.179	0.198
		(0.22)	(0.24)	(0.24)	(0.25)
friend				0.233	0.256
				(0.21)	(0.21)
1.clubmate's now in serious relationship*friend				0.287	0.272
				(0.47)	(0.43)
own age					0.544
					(0.52)
own highest high school level					1.247
					(1.43)
own hh relative income					0.892
					(0.78)
Obs	658	650	650	650	650
Pseudo R-squared	0.003	0.133	0.133	0.136	0.200
Clustered standard errors	NO	NO	YES	YES	YES
Friend	NO	NO	NO	YES	YES
Controls	NO	NO	NO	NO	YES

Note: Dependent variable is own current relationship status (1=currently in a relationship, 0=otherwise). Standard errors are given in parentheses. In case of clustering, standard errors are clustered at the individual level. *** indicates significance at the 0.01 level; ** indicates significance at the 0.05 level; * indicates significance at the 0.10 level. The first digit of the coefficient name indicates the category number. Outliers are included.

Table 4.C18: Ordered logistic regression of own current number of bed partners on categories of own lagged bed partners and clubmate's current and lagged bed partners

	ologit1	ologit2	ologit3	ologit4	ologit5
1.clubmate's bed partners	-0.438	-0.166	-0.166	-0.325	-0.416
	(0.30)	(0.36)	(0.36)	(0.44)	(0.42)
2.clubmate's bed partners	-0.545*	-0.534	-0.534**	-0.723**	-0.764**
	(0.31)	(0.38)	(0.22)	(0.30)	(0.31)
3.clubmate's bed partners	0.005	0.375	0.375	0.158	-0.084
	(0.34)	(0.45)	(0.79)	(0.82)	(0.77)
4.clubmate's bed partners	0.111	0.191	0.191	-0.223	-0.368
	(0.33)	(0.57)	(0.66)	(0.65)	(0.64)
own lag bed partners		3.107***	3.107***	3.107***	3.176***
		(0.19)	(0.68)	(0.68)	(0.70)
clubmate's lag bed partners		0.011	0.011	-0.005	0.043
		(0.15)	(0.15)	(0.14)	(0.13)
friend				-0.492	-0.526
				(0.57)	(0.62)
1.clubmate's bed partners*friend				0.467	0.466
				(0.66)	(0.68)
2.clubmate's bed partners*friend				0.575	0.647
				(0.69)	(0.73)
3.clubmate's bed partners*friend				0.842	0.710
				(0.83)	(0.76)
4.clubmate's bed partners*friend				1.304**	1.331**
				(0.63)	(0.63)
own age					-0.024
					(0.42)
own highest high school level					-1.213*
					(0.71)
own hh relative income					0.036
					(0.58)
Obs	658	632	632	632	632
Pseudo R-squared	0.007	0.319	0.319	0.321	0.340
Clustered standard errors	NO	NO	YES	YES	YES
Friend	NO	NO	NO	YES	YES
Controls	NO	NO	NO	NO	YES

Note: Dependent variable is own current number of bed partners (1=0 bed partners, 2=1-3, 3=4-6, 4=7-9, 5=10 or more). Standard errors are given in parentheses. In case of clustering, standard errors are clustered at the individual level. *** indicates significance at the 0.01 level; ** indicates significance at the 0.05 level; * indicates significance at the 0.10 level. The first digit of the coefficient name indicates the category number. Outliers are included.

Table 4.C19: Ordered logistic regression of own current grade on categories of own lagged grade and clubmate's current and lagged grade

	ologit1	ologit2	ologit3	ologit4	ologit5
6.clubmate's grade	0.337	0.154	0.154	0.185	0.312
	(0.32)	(0.33)	(0.33)	(0.37)	(0.37)
7.clubmate's grade	0.036	-0.141	-0.141	-0.029	0.079
	(0.30)	(0.31)	(0.27)	(0.30)	(0.30)
8.clubmate's grade	0.194	-0.082	-0.082	0.414	0.602
	(0.37)	(0.38)	(0.38)	(0.46)	(0.45)
own lag grade		0.492***	0.492	0.484	0.492
		(0.12)	(0.42)	(0.42)	(0.47)
clubmate's lag grade		0.169	0.169	0.178	0.207*
		(0.12)	(0.11)	(0.11)	(0.11)
friend				0.824	0.758
				(0.54)	(0.55)
6.clubmate's grade*friend				-0.432	-0.449
				(0.70)	(0.71)
7.clubmate's grade*friend				-0.678	-0.704
				(0.63)	(0.63)
8.clubmate's grade*friend				-1.699**	-1.739**
				(0.66)	(0.68)
own age					0.587*
					(0.36)
own highest high school level					0.516
					(0.70)
own hh relative income					1.127*
					(0.66)
Obs	616	616	616	616	616
Pseudo R-squared	0.002	0.016	0.016	0.021	0.080
Clustered standard errors	NO	NO	YES	YES	YES
Friend	NO	NO	NO	YES	YES
Controls	NO	NO	NO	NO	YES

Note: Dependent variable is own current average grade (1-10). Standard errors are given in parentheses. In case of clustering, standard errors are clustered at the individual level. *** indicates significance at the 0.01 level; ** indicates significance at the 0.05 level; * indicates significance at the 0.10 level. The first digit of the coefficient name indicates the category number. Outliers are included.

Chapter 5

The Smart, the Lucky and the Greedy: Income Inequality and Trust in the Lab

Joint work with David Smerdon

> "I thought the roulette part was a bit unfair, it was just random."
>
> – Anonymous subject

5.1 Introduction

In a recent speech, US President Barack Obama declared a war on socioeconomic inequality, calling it "the defining challenge of our time."[1] Obama stressed the negative effects of inequality on social phenomena, particularly trust. While the disadvantages of income inequality to various economic outcomes have been explored for some time (Eccles, 1966; Ostry et al., 2014), the interest in the linkages between inequality and 'softer' factors is relatively new. These include such indicators as social cohesion, self-esteem, and trust.

[1] http://www.whitehouse.gov/the-press-office/2013/12/04/remarks-president-economic-mobility, accessed 15 June 2014

Trust "is an important lubricant in a social system" (Arrow, 1974, p.23), offering a range of valuable economic and social benefits to a society (Fukuyama, 1995). A large body of empirical research has found evidence to suggest a negative correlation between income inequality and trust (Pickett and Wilkinson, 2009). Why should the income distribution mechanism matter to trust? To answer this question, it is worth first considering the definition of trust as given by Gambetta (2000): "When we say we trust someone or that someone is trustworthy, we implicitly mean that the probability that he will perform an action that is beneficial or at least not detrimental to us is high enough for us to consider engaging in some form of cooperation with him" [p. 217]. This probability depends on the value the trustee places on non-pecuniary factors - notably fairness, inequity and reciprocity considerations. Social norms with regards to these factors may in turn be influenced by the class structure. Several authors have consequently highlighted social norms as driving the development of trust within a society (Kramer, 1996; Krueger et al., 2008). This is related to the "accountability principle", whereby individuals are said to display a more egalitarian approach to fairness when a social surplus arises from actions without intentionality (Cappelen et al., 2007; Konow, 1996, 2000). Evidence from the dictator and ultimatum games also supports the claim that individuals who have earned their position of deciding power are less inclined to offer a share of a surplus to another (Cherry et al., 2002; Hoffman et al., 1994). It follows that the procedure leading to a state of income inequality may impact how the level of inequality affects societal trust.

Empirical studies on trust and inequality generally use survey measures of trust, and apply instrumental variables to address potential endogeneity issues (Alesina and La Ferrara, 2002; Gustavsson and Jordahl, 2008; Leigh, 2006a,b). Endogeneity could stem from a reverse causal relationship between trust and inequality, as well as omitted variables such as cultural factors. For example, cultural norms with regards to sharing could affect both trust and inequality. There are three main issues with such studies. Firstly, there are concerns with regard to the exogeneity of the instruments commonly chosen (Jordahl, 2008). Combined with likely measurement errors in inequality data, this creates some

doubt over the validity of empirical results. Secondly, due to the self-reported and subjective nature, survey-based trust measures are prone to respondent issues such as hypothetical bias, idealised personal bias, and lack of incentive compatibility. Ciriolo (2007) discusses these biases in his argument for the use of laboratory experiments in estimating relationships pertaining to trust. Finally, a common criticism of empirical studies into the relationship between trust and inequality is that this research does not acknowledge social, historical, and cultural differences between countries (Snowdon, 2010). With regard to the closely related body of literature on inequity aversion and fairness, many researchers claim that perceptions of social values highly depend on context (e.g. Shiller et al., 1991).

This paper compares the effect of different degrees of income inequality on generalised trust (that is, trust of strangers). Generalised trust indicates whether a person believes that individuals in her group (e.g. country) can generally be trusted. By studying this relationship in the laboratory, we tackle endogeneity concerns pertaining empirical studies. Also, we use an experimental game to measure trust rather than relying on self-reported, unincentivised, measures. Our main contribution is to test whether the relationship between inequality and trust is robust to changes in the *income distribution mechanism*. This is a relevant consideration, given the importance of contextual heterogeneity between societies. For example, income class divisions can be a signal of past meritorious or, conversely, unscrupulous behaviour such as corruption. Divisions can even be a reflection of a class structure where one's position is random such as in a caste system. Specifically we explore the differences in trust when income position within a society is allocated randomly, based on merit, or due to greed. We find that the income distribution mechanism matters: higher inequality lowers trust levels when income is distributed randomly, but has no effect on trust when income is distributed on the basis of merit or greed.

Trust and inequality in the lab

While the empirical literature generally uses survey measures of trust, the most common measurement tool in the lab is the trust game of Berg et al. (1995). (Johnson and Mislin, 2012) show that the

results from survey measures are positively correlated with experimentally measured trust. Sapienza et al. (2013) argue that the subject's expectation about the actions of others in laboratory trust games is a better measure for generalised trust because it is not contaminated by other-regarding and risk preferences. Consequently, our design accounts for both expectations and observed actions. Moreover, we can further close the distance between these measures by ensuring the trust game is played between anonymous partners, as this better replicates the environment for generalised trust.[2]

There have been a handful of papers in recent years that feature anonymous trust games treated with heterogeneous endowments, although most have found no effect of this treatment (Anderson et al., 2006; Brülhart and Usunier, 2012). Small but significant effects have been recorded when subjects know the endowment of their partner (Hargreaves Heap et al., 2013; Smith, 2011). However, the results seem to go in opposite directions: Smith (2011) reports that low income subjects send more in the trust game to a high income subject than to a fellow low income subject, while Hargreaves Heap et al. (2013) and Greiner et al. (2012) find a damaging effect of income inequality on trust. Finally, Xiao and Bicchieri (2010) show that subjects are less likely to reciprocate in the trust game if it increases earnings inequality (i.e. when paired with a trustor of higher endowment).

Smith (2011) highlights that future studies on trust and inequality should look beyond the random distribution of endowments, adding: "Past experiments suggest that endowment origin affects behaviour" [p.56]. In line with this view, we study the impact of nonrandom income distribution mechanisms, which has not been done before in the literature concerning trust and inequality. In particular, our research investigates merit- and greed-based income distributions. Another novel feature of our study is the elicitation of expectations in the trust game, following Sapienza et al. (2013). This measure is, unlike the Sender behaviour, not influenced by other-regarding and risk preferences, and is therefore a more adequate measure of trust.

[2]For example, the World Values Survey, the predominantly quoted survey measure of trust, asks respondents: "Generally speaking, would you say that most people can be trusted, or that you can't be too careful in dealing with people"

Trust, inequality and income distribution mechanisms

Most closely related to our paper is Ku and Salmon (2013). To our knowledge, this is the only preceding study to experimentally investigate the influence of the income distribution mechanism on attitudes to inequality. In their experiment, subjects play an investment game in pairs of 'rich' and 'poor' subjects. Income positions are determined on the basis of a random, meritocratic or greedy criterion, as is the case in our experiment. In a fourth treatment, they use an arbitrary criterion to determine income class. They find that individuals in the disadvantaged class make investment decisions with lower efficiency when nonrandom criteria are used to distribute endowments. Their research may suggest that individuals are less tolerant of inequality when the society's income distribution mechanism is contingent on the intentioned actions of its members. In addition, Ku and Salmon (2013) show that subjects transfer more to members of their own group.

This latter finding illustrates how the response to income inequality plausibly depends on the strength of ingroup favouring (see also Lei and Vesely, 2010). This is linked to the claim by social identity theory that the existence of stereotypes is based on salient group membership, whereby individuals have a higher propensity to trust or recipocrate ingroup than outgroup members (Chen and Li, 2009; Platow et al., 2012; Tanis and Postmes, 2005). Consider an individual who lives in a society in which its subjects are placed into fixed income classes at random - imagine, perhaps, a prevalent caste structure. Her attitude to a fair sharing of a (randomly-produced) surplus could conceivably differ from that in a society in which membership of the higher income class results from meritorious effort, or alternatively from past 'greedy' actions at the expense of others, such as corruption. A culture of "us versus them" could more likely be fostered if it is believed that the rich achieved their position in an undeserving or reprehensible manner. The mechanism of the distribution can thus be responsible for the strength of an ingroup/outgroup environment. Uslaner (2008) succinctly summarises this connection between generalised trust and the common bonds within classes: "If we believe that we have a shared fate with others, and especially people who are different to ourselves, then gross inequalities in wealth

and status will seem to violate norms of fairness" [p.49].

In our laboratory design, subjects are first assigned to either a small, high-income class or a larger, low-income class, following a merit-based, greed-based or random allocation. A further treatment variable is the degree of inequality. Subjects then play the trust game against anonymous partners, including the elicitation of expectations with regards to the trustworthiness of their opponent. Our main findings can be summarised thusly: Higher income inequality impacts trust in a society in which income classes are determined randomly. When the income distribution mechanism is based on either merit or greed, however, we cannot conclude that changes in income inequality affect trust within the group. Our findings are robust against selection effects. Also, we show that expectations, rather than sending behaviour, are correlated to survey measures of trust. We suggest in the discussion of the results that our results may be driven by the influence of the distribution mechanism on ingroup/outgroup effects.

The remainder of the paper is typically structured. Section 5.2 details our experimental design. Section 5.3 presents the results, Section 5.4 includes robustness checks, and Section 5.5 discusses possible channels for the effects. Section 5.6 concludes.

5.2 Experimental Design

Our experiment has a 2x3 design. In each session, subjects are either placed in a high income inequality (High) treatment or a low income inequality (Low) treatment. Next, our treatments differ in terms of 'mechanism': endowments are distributed randomly (Random), based on merit (Merit) or based on greed (Greed). We label the resulting six treatments: RandomHigh, RandomLow, MeritHigh, MeritLow, GreedHigh, and GreedLow. Table 5.1 provides an overview of the treatments. The instructions of the experiment are included in Appendix 5.A.

Our experiment has two stages, of which the second stage is divided in two periods. In the first stage, all subjects play three different tasks. In the 'roulette task', subjects pick a number on a roulette

wheel with 36 slots. In the 'calculation task', subjects have four minutes to solve as many calculation problems as they can. They are asked to find the highest number in each of two matrices and compute their sum. In the 'decider task', subjects play a variation of the standard dictator game in which the amount allocated to the (anonymous) partner is doubled. Before the start of the first stage, subjects are informed that there will be a second stage, and that their performance in the first stage will affect the second stage. They are, however, not notified what this effect entails.

The second stage is divided into two periods with a similar design. In both periods, subjects play a trust game with endowments that are assigned based on their performance in the first stage. One of the three tasks from the first stage - the 'allocation task' - is used to assign subjects to their income group (with its respective endowments). Subjects are informed at the beginning of the second stage which task is selected. In the Random treatments, endowments are based on performance in the roulette task: 25% of the subjects who picked a number closest to the winning number will receive the 'high' endowment. Correspondingly, in the Merit treatment the high endowment is received by the 25% of the subjects who solved the most calculations correctly in the calculation task, and in the Greed treatments by the 25% who took the most money in the decider task.[3]

At the beginning of the first period of the second stage, subjects learn the distribution of endowments for this period. Endowments are divided into two classes: high and low. 25% of the subjects receive the high endowment and 75% receive the low endowment. The low endowment is the same in the High and Low inequality treatments, but the high endowment differs across treatments. In all treatments, subjects in the low income group receive as their endowment 160 experimental tokens per period (one token equals one euro cent). Subjects in the high income group receive an endowment of 300 tokens in the High treatments or 180 tokens in the Low treatments.[4]

[3] In case of ties, we randomly allocate tied subjects to the high- and low-income group.

[4] We intentionally do not have a treatment with complete equality for two reasons. First, our distribution mechanisms would become irrelevant in case of full equality. Second, we want to test whether it is indeed *higher* inequality that impedes trust. Current studies investigate whether trust is lower in (any) inequality than in full equality. In these studies, the inequality effect could be caused merely by leaving a state of full inequality.

Having been informed about the income distribution, subjects play the trust game from Berg et al. (1995) with the endowment they received in that period. As was explained in the Introduction, we choose this game, because - when played with an anonymous opponent - it approximates the concept of generalised trust. First, the 'Sender' decides how much of her endowment (in multiples of twenty tokens) to pass on to the 'Receiver'. The roles of Sender and Receiver represent those of the trustor and trustee, respectively. The amount sent by the Sender is tripled upon receipt. Next, the Receiver decides how much money to return. Note that this amount will not be multiplied. The amount sent by the Sender is a measure of trust; the amount returned by the Receiver is a measure of trustworthiness. Each participant plays both as Sender and Receiver. We use a full strategy method for the Receiver's decision, recording for each possible amount sent by the Sender what the Receiver would return.

In addition to asking for the Sender's and Receiver's responses, we elicit the Sender's expectations about the Receiver's behaviour after the Sender has decided how much money to send. For each possible amount sent by the Sender, the Sender indicates how much she expects to receive back from the Receiver. This 'expected trustworthiness' is an alternative measure for trust, and is arguably more accurate since it is not contaminated by risk and other-regarding preferences (Sapienza et al., 2013).

After finishing the trust game in the first period of the second stage, subjects are informed about the distribution of endowments for the second period. Subjects remain in the same income group: those who received the low (high) endowment in the first period, again receive the low (high) endowment in the second period. The low endowment remains the same, but the high endowment changes: in the High (Low) treatment, the high endowment in the second period equals 180 (300) tokens. Hence, in the High treatments, subjects move from high to low inequality, and vice versa for the Low treatments. There is no feedback in between the periods. After learning the new income distribution, subjects play the trust game once more, with the endowments they received in the second period.

Of the tasks in the first stage, only the allocation task is paid out. Subjects learn after the first stage which task is selected to be the allocation task, and are not told their earnings until the end

of the experiment. The payoffs from the three tasks are structured so as to be appropriately equal in expectation (around 300 tokens). In the second stage, all three decisions - including the Sender's expectations - in both periods are incentivised. Subjects are matched with a different, anonymous subject for each decision. The payoff for the expectations of the Receiver's behaviour is based on a randomly selected hypothetical amount sent. Subjects receive 100 tokens if their guess was within 10% of the amount returned by their matched partner for this amount.

At the end of the experiment, we measure subjects' risk aversion by using the lottery task of Holt and Laury (2002). We intentionally choose to describe the lotteries to the subjects in terms of euros, not tokens, so subjects realise that it is not part of the main experiment. We also conduct an exit survey, including questions about demographics, fairness, trust and inequality.

The experiment was run in April 2014 in the CREED laboratory of the University of Amsterdam. The duration of each session was roughly an hour. The participants were all recruited from the CREED database, through an email notification. Most of our subjects - 98% - are students. The experiment was programmed in zTree (Fischbacher, 2007). In total, 240 subjects participated in the experiment - 40 per treatment, broken into two sessions of 20 subjects. The average earnings were 15.34 euro, which included a show-up fee of 3.00 euro.

Table 5.2 shows the descriptive statistics for our full sample and the separate treatments. All differences between the treatments are statistically insignificant (Bonferroni multiple-comparison test), except for one: the number of correct answers in the calculation task is statistically different between MeritHigh and RandomLow (p-value 0.052). As we never compare these two treatments in our analysis, this significant difference does not influence our results.

Table 5.1: Treatments

	Treatment	Allocation task	Inequality in period 1	Inequality in period 2
1.	RandomHigh	Roulette	High	Low
2.	RandomLow	Roulette	Low	High
3.	MeritHigh	Calculation	High	Low
4.	MeritLow	Calculation	Low	High
5.	GreedHigh	Decider	High	Low
6.	GreedLow	Decider	Low	High

Table 5.2: Descriptive Statistics

	All	RandomHigh	RandomLow	MeritHigh	MeritLow	GreedHigh	GreedLow
Earnings in euro	15.34	15.43	15.82	16.07	15.22	14.58	14.89
	(3.43)	(3.68)	(4.04)	(3.40)	(2.47)	(2.96)	(3.79)
Correct answers in calculation task	10.01	10.20	9.07	11.00	10.18	9.85	9.78
	(2.94)	(2.89)	(2.86)	(3.17)	(2.86)	(2.68)	(3.00)
Tokens kept in decider task	424.66	429.80	411.48	439.95	414.73	419.57	432.43
	(92.82)	(97.02)	(115.49)	(72.75)	(78.72)	(114.33)	(69.63)
Female	0.54	0.45	0.55	0.57	0.62	0.50	0.55
	(0.50)	(0.50)	(0.50)	(0.50)	(0.49)	(0.51)	(0.50)
Age	22.34	22.75	22.75	21.75	22.62	21.75	22.43
	(2.82)	(2.76)	(3.25)	(2.36)	(3.02)	(3.13)	(2.24)
Risk aversion	5.92	5.85	5.75	5.75	6.10	6.15	5.92
	(1.63)	(2.15)	(1.53)	(1.37)	(1.52)	(1.63)	(1.54)
Observations	240	40	40	40	40	40	40

Note: The table reports the mean of selected variables. Standard deviations are given in parentheses. Risk aversion is measured by the number of safe choices in the lottery from Holt and Laury (2002).

5.3 Results

In this section we discuss the results with regards to three main outcome variables. The first is 'trust', the amount of money sent by the Sender in the trust game. The second is 'expectations', the amount the Sender expects to receive back for a hypothetical amount sent. The third is 'trustworthiness', which is the amount the Receiver would return for a hypothetical amount received. For expectations and trustworthiness we have multiple observations per subject, for each hypothetical amount sent or received. Descriptive statistics of the outcome variables are included in Appendix Tables 5.B1, 5.B2,

5.B3, and 5.B4.

We focus our discussion of the results on the low-income[5] group in period 1. Subjects in the high-income group have a larger income in high inequality than in low inequality. It is therefore not possible in our design to disentangle the effect of inequality from the wealth effect. In addition, our sample of high-income subjects is - by design - much smaller than the sample of low-income subjects. Furthermore, we focus on the first period to avoid priming effects from activating 'consequential thinking' (Kugler et al., 2009), meaning that subjects are aware of the consequences of their actions. In any case, as we will show later, the inequality effect on trust is highly persistent.

This section first looks into the impact of inequality on trust across our three mechanisms. Next, we discuss whether trust levels vary across mechanisms. We then briefly look at the results for the second period and the high-income subjects. In the last subsection, we discuss the salience of income differences and perception of fairness across the mechanisms.

5.3.1 Testing inequality

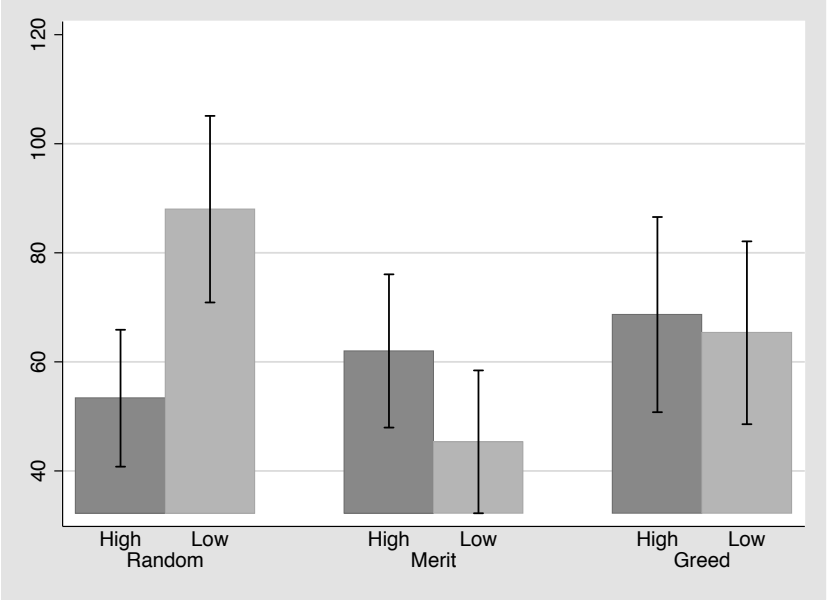
Figure 5.1 shows trust across the six treatments for the low-income subjects. The nonparametric results are included in Table 5.3.[6] Trust is significantly higher in high than in low inequality when income is distributed randomly, but this is not the case when income is distributed based on merit or greed. In fact, we find a significantly lower level of trust in MeritLow as compared to MeritHigh. Ku and Salmon (2013) find a similar result and suggest that it might be caused because of an emotional response after 'losing' in the effort task. We will show below that this difference is not found in expectations and trustworthiness, which are arguably less influenced by emotions.

Figure 5.2 depicts the expectations and trustworthiness for the low-income subjects by mechanism

[5] We use 'endowment' and 'income' interchangeably.

[6] We use the Mann-Whitney rank sum test, which tests the null hypothesis that two independent samples come from the same distribution. We use this test as it is appropriate for a between-subjects design like ours and does not require any assumptions with regards to the distribution of the variable of interest.

Figure 5.1: Trust for low-income group



Note: The amount sent in period 1 for low-income subjects in high and low income inequality, according to mechanism. Error bars indicate 10% confidence intervals. Amounts are in tokens.

and Table 5.4 includes the nonparametric test results for these variables. We find no significant differences between low and high inequality when the income distribution is based on merit or greed. In contrast, expectations are higher overall in the RandomLow treatment than in the RandomHigh treatment. Trustworthiness is also higher in RandomLow than in RandomHigh, and these differences are significant for received amounts of 80 tokens or more.

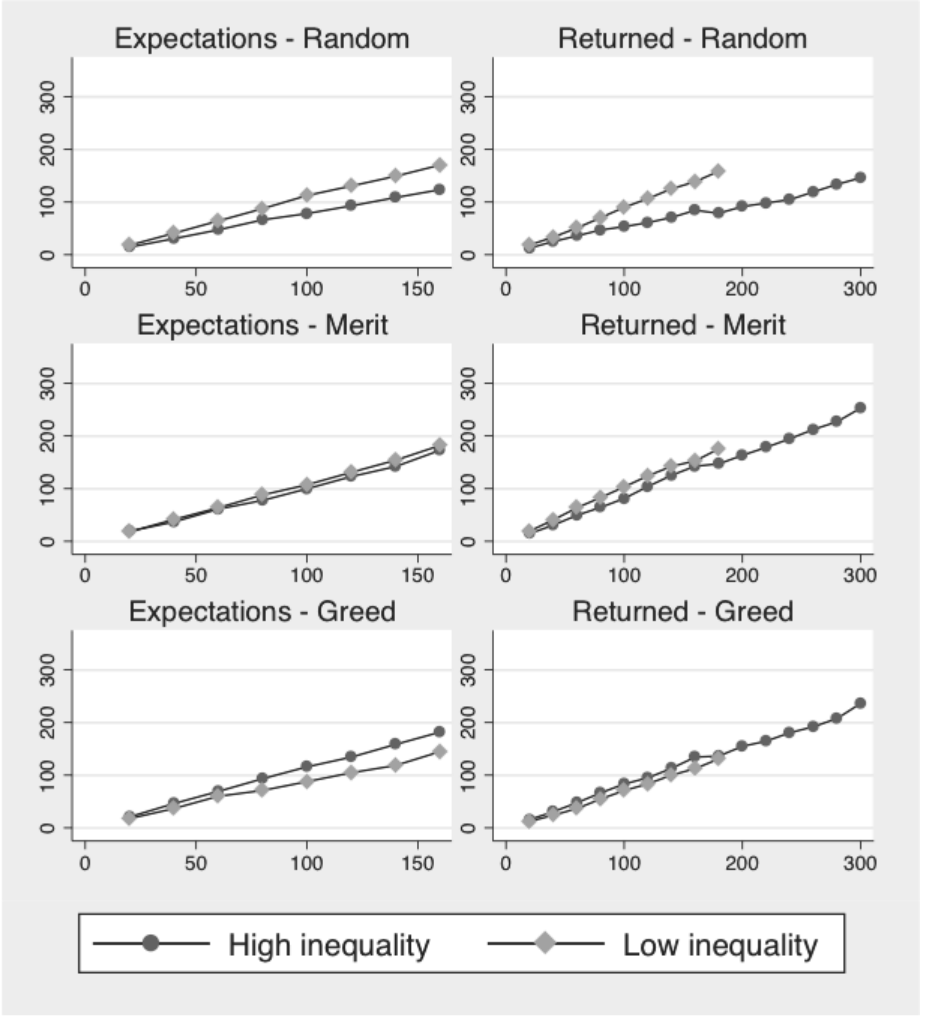
In short, we find an inequality effect for trust, expectations and trustworthiness, but only when income is distributed randomly. When income is distributed based on greed or merit, inequality generally does not have an effect.

Table 5.3: Trust differences between treatments for low-income group

				Difference
Testing inequality	RandomHigh	-	RandomLow	-34.67**
	MeritHigh	-	MeritLow	16.67*
	GreedHigh	-	GreedLow	3.33
Testing mechanism	RandomHigh	-	MeritHigh	-8.67
(high inequality)	RandomHigh	-	GreedHigh	-15.33
	MeritHigh	-	GreedHigh	-6.67
Testing mechanism	RandomLow	-	MeritLow	42.67***
(low inequality)	RandomLow	-	GreedLow	22.67
	MeritLow	-	GreedLow	-20.00

Note: The table reports differences in trust between the treatment reported in the second column and the treatment reported in the third column for low-income subjects. The treatments either have the same mechanism ('Testing inequality') or the same level of inequality ('Testing mechanism'). Amounts are in tokens. ***, ** and * indicates significance at the 1%, 5% and 10% respectively in the Mann-Whitney rank sum test.

Figure 5.2: Expectations and trustworthiness for low-income group by Mechanism



Note: The figure shows the expectations (left column) and trustworthiness (right column) for the Random, Merit and Greed treatment respectively for low-income subjects. In the left column, the x-axis indicates the hypothetical amount sent (ranging from 0 to 160 for low-income subjects) and the y-axis indicates the expected amount returned. In the right column, the x-axis indicates the hypothetical amount received before tripling (ranging from 0 to 300) and the y-axis indicates the amount returned. Amounts are in tokens.

Table 5.4: P-values for non-parametric tests for inequality and mechanism for low-income group

Panel A: Expectations	*Null hypothesis*																		
Amount sent				20	40	60	80	100	120	140	160								
Testing inequality	RandomHigh	-	RandomLow	-3.60	-10.27*	-16.93*	-21.27	-34.27**	-37.60*	-40.60*	-46.77*								
	MeritHigh	-	MeritLow	1.03	-4.63	-2.47	-10.30	-7.63	-8.13	-12.00	-9.17								
	GreedHigh	-	GreedLow	3.17	9.50	9.17	22.33	28.00	29.67	39.97	37.17								
Testing mechanism	RandomHigh	-	MeritHigh	-4.63	-6.27	-13.97	-11.47	-21.47	-29.97	-33.43	-49.27								
(high inequality)	RandomHigh	-	GreedHigh	-6.43	-15.60*	-21.93	-27.10	-37.60	-41.43	-49.90	-58.27								
	MeritHigh	-	GreedHigh	-1.80	-9.33	-7.97	-15.63	-16.13	-11.47	-16.47	-9.00								
Testing mechanism	RandomLow	-	MeritLow	0.00	-0.63	0.50	-0.50	5.17	-0.50	-4.83	-11.67								
(low inequality)	RandomLow	-	GreedLow	0.33	4.17	4.17	16.50	24.67	25.83	30.67	25.67								
	MeritLow	-	GreedLow	0.33	4.80	3.67	17.00	19.50	26.33	35.50	37.33								
Panel B: Trustworthiness	*Null hypothesis*																		
Amount received				20	40	60	80	100	120	140	160	180	200	220	240	260	280	300	
Testing inequality	RandomHigh	-	RandomLow	-5.83	-8.50	-15.00	-23.17*	-36.83**	-45.50**	-54.43**	-54.27*	-79.90**							
	MeritHigh	-	MeritLow	-4.83	-10.00	-14.33	-18.67	-22.00	-20.33	-17.00	-9.93	-28.67							
	GreedHigh	-	GreedLow	3.33	6.67	11.00	11.67	12.50	11.50	12.83	21.83	5.33							
Testing mechanism	RandomHigh	-	MeritHigh	-1.83	-5.67	-13.50	-17.33	-28.00	-43.00**	-54.60**	-58.27*	-68.23*	-71.87*	-80.17*	-90.13*	-93.07*	-93.33*	-107.00*	
(high inequality)	RandomHigh	-	GreedHigh	-3.00	-5.67	-11.83	-19.33	-30.00	-34.00	-42.27	-49.93	-57.57	-63.37	-66.50	-76.13	-72.73	-74.00	-90.33	
	MeritHigh	-	GreedHigh	-1.17	0.00	1.67	-2.00	-2.00	9.00	12.33	8.33	10.67	8.50	13.67	14.00	20.33	19.33	16.67	
Testing mechanism	RandomLow	-	MeritLow	-0.83	-7.17	-12.83	-12.83	-13.17	-17.83	-17.17	-13.93	-17.00							
(low inequality)	RandomLow	-	GreedLow	6.17*	9.50 *	14.17 *	15.50 *	19.33	23.00	25.00	26.17	27.67							
	MeritLow	-	GreedLow	7.00	16.67**	27.00 **	28.33 *	32.50*	40.83 *	42.17	40.10	44.67							

Note: The table reports differences between the expectations and trustworthiness (measured in tokens) in the treatment reported in the second column and the treatment reported in the third column for low-income subjects. The treatments either have the same mechanism ('Testing inequality') or the same level of inequality ('Testing mechanism'). ***, ** and * indicates significance at the 1%, 5% and 10% respectively in the Mann-Whitney rank sum test.

5.3.2 Testing mechanism

Next, we compare the levels of the outcome variables across the three mechanisms. Figure 5.1 and Table 5.3 show that trust is relatively low in RandomHigh and relatively high in RandomLow compared to the treatments. Trust levels in GreedLow and GreedHigh are similar and are situated in between the levels of RandomHigh and RandomLow. Trust is low in both MeritHigh and MeritLow. See Table 5.4 for nonparametric test results concerning expectations and trustworthiness.[7] In high inequality the lowest expectations and trustworthiness are reported for the Random treatment, while these outcome variables are the lowest in the Greed treatment when inequality is low. The differences in expectations are not significant. Trustworthiness, however, is significantly higher in both Merit treatments. In short, we find mixed results when we compare the levels of trust, expectations, and trustworthiness across mechanisms in the low-income group.

5.3.3 Second period

In the second period of our experiment, we keep the mechanism the same, but switch the size of the income differences. A treatment with low inequality in the first period switches to high inequality in the second period, and vice versa. These results should be compared with some caution, as there is experimental evidence to suggest a priming effect for 'consequentialist thinking' in trust games after expectations have been elicited (Kugler et al., 2009). Additionally, we need to take into account order effects. We therefore regress second period variables on their lagged variable and dummies for the treatments. Table 5.5 shows that first period choices largely determine second period decisions. All lagged variables are highly significant and treatment dummies are never significant. R-squared values indicate that between 50% and 72% of the variance is explained. We infer from this result that second period observations are not informative of the effect of inequality. It is therefore not surprising that our

[7] For convenience, Figure 5.C1 shows expectations and trustworthiness by income distribution rather than by mechanism.

findings for the second period are similar to the first period. If we repeat our empirical analysis for the second period, we find that trust, expectations, and trustworthiness remains lower in RandomHigh than in RandomLow; we generally find no differences between high and low inequality in the Merit and Greed treatments.[8]

5.3.4 High-income group

Bearing the potential wealth effects and small sample size in mind, we look at the inequality effect for high-income subjects. Figure 5.3 shows that high-income subjects do not necessarily send more money when they have a higher income, i.e. when inequality is higher. The difference between high and low inequality is only significant when income is distributed randomly (see Table 5.6 for nonparametric test results). Hence, we find again that inequality only affects trust for a random income distribution, but - unlike the low-income subjects - high-income subjects trust more in high inequality. In line with the results for the low-income group, the high-income group has higher expectations and trustworthiness in RandomLow than in RandomHigh (Figure 5.4 and Table 5.7). Expectations are not significantly different between high and low inequality in the Merit and Greed treatments. Trustworthiness is not different between MeritHigh and MeritLow, but it is significantly higher in GreedHigh than GreedLow for most amounts received.

What about the level of trust across mechanisms? Trust is higher in case of random distributions than in greed- or merit-based distributions both in high and low inequality (see Figure 5.1 and Table 5.6). If we look at expectations and trustworthiness, we again find the highest levels when income is distributed randomly (Table 5.7).[9] Interestingly, Figure 5.4 shows that trustworthiness is close to zero for all possible amounts received in the GreedLow treatment.

[8] See Figures 5.C3 and 5.C4, and Tables 5.B5 and 5.B6.

[9] For convenience, Figure 5.C2 shows expectations and trustworthiness by income distribution rather than by mechanism.

Table 5.5: Regressions for low-income group, second period

	Trust	Expectations	Trustworthiness
Constant	12.45*	3.00	4.48
	(7.30)	(3.91)	(5.28)
RandomLow	3.79	1.08	-1.03
	(9.66)	(6.81)	(9.59)
MeritHigh	8.78	-4.07	-0.51
	(9.49)	(6.14)	(7.63)
MeritLow	6.10	9.06	-0.10
	(9.49)	(8.80)	(8.19)
GreedHigh	10.25	-11.84	6.15
	(9.51)	(10.61)	(9.69)
GreedLow	8.52	3.27	-13.12
	(9.50)	(9.84)	(12.25)
Trust	0.68***		
	(0.05)		
Expectations		0.80***	
		(0.07)	
Trustworthiness			0.80***
			(0.07)
Obs	180	1620	1620
R-squared	0.50	0.70	0.72
Sent/received amount fixed effects	No	Yes	Yes
Clustered standard errors	No	Yes	Yes
P-values for Wald test			
Testing inequality			
MeritHigh=MeritLow	0.78	0.16	0.95
GreedHigh=GreedLow	0.85	0.28	0.18
Testing mechanism			
RandomHigh=MeritHigh	0.36	0.51	0.95
RandomHigh=GreedHigh	0.28	0.27	0.53
MeritHigh=GreedHigh	0.88	0.50	0.43
RandomLow=MeritLow	0.81	0.42	0.91
RandomLow=GreedLow	0.62	0.84	0.41
MeritLow=GreedLow	0.80	0.64	0.37

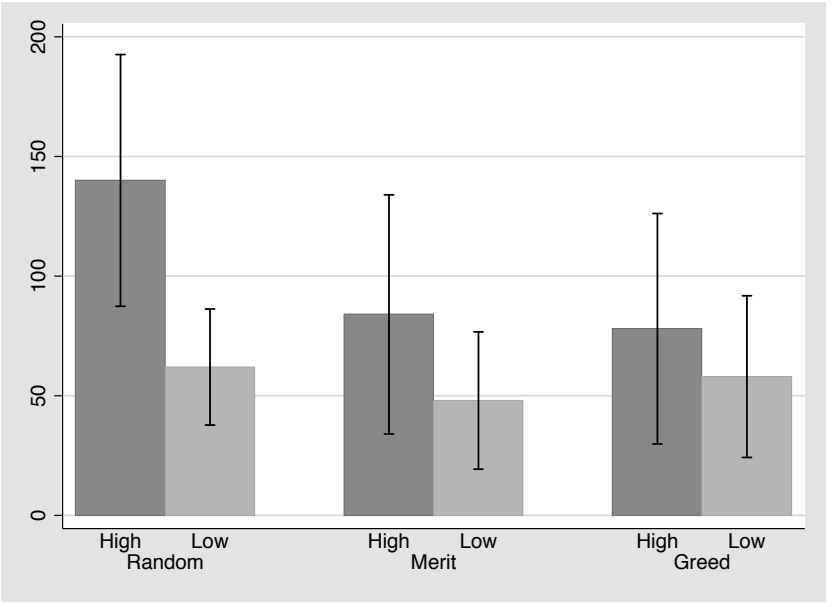
Note: OLS regressions for low-income subjects of second period trust, expectations and trustworthiness (measured in tokens) on their lagged variables and treatments. ***, ** and * indicates significance at the 1%, 5% and 10% respectively. Standard errors are given in parentheses.

Table 5.6: Trust differences between treatments for high-income group

				Difference
Testing inequality	RandomHigh	-	RandomLow	78.00*
	MeritHigh	-	MeritLow	36.00
	GreedHigh	-	GreedLow	20.00
Testing mechanism	RandomHigh	-	MeritHigh	56.00
(high inequality)	RandomHigh	-	GreedHigh	62.00
	MeritHigh	-	GreedHigh	6.00
Testing mechanism	RandomLow	-	MeritLow	14.00
(low inequality)	RandomLow	-	GreedLow	4.00
	MeritLow	-	GreedLow	-10.00

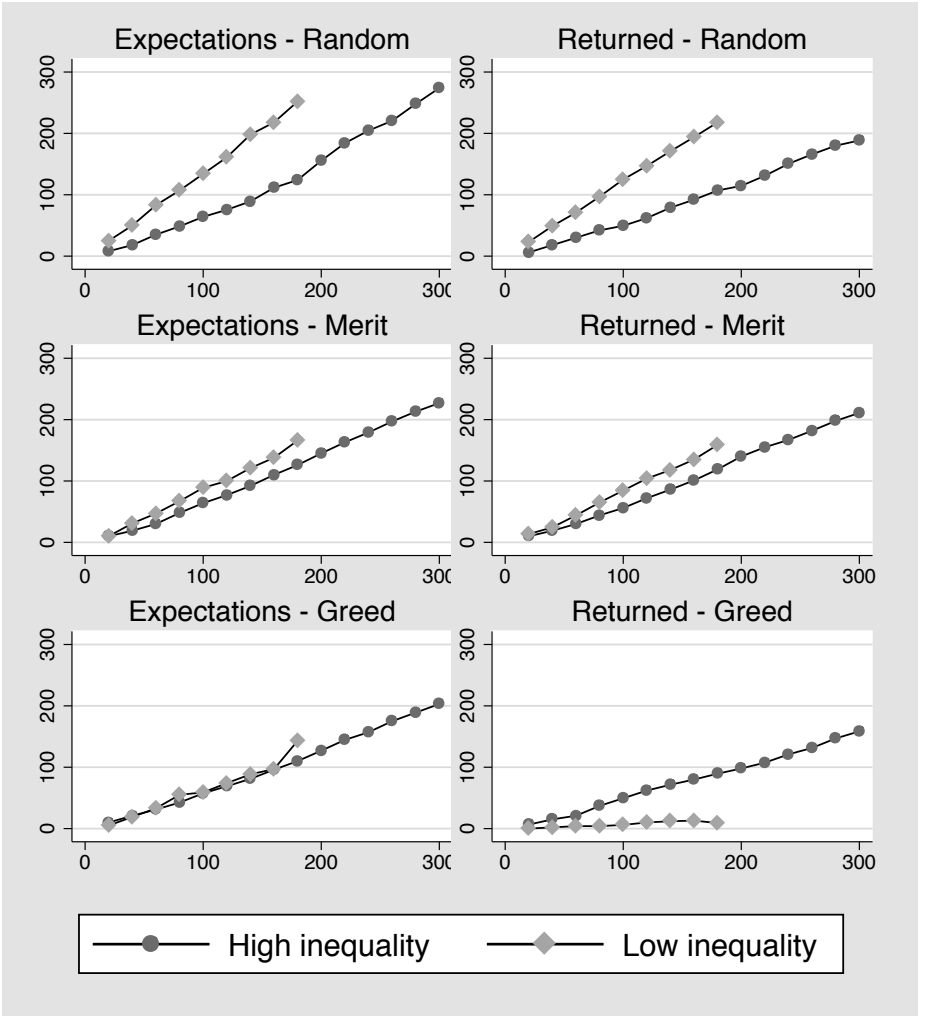
Note: The table reports differences in trust (measured in tokens) between the treatment reported in the second column and the treatment reported in the third column for high-income subjects. The treatments either have the same mechanism ('Testing inequality') or the same level of inequality ('Testing mechanism'). The variable Trust is the amount sent in the trust game. ***, ** and * indicates significance at the 1%, 5% and 10% respectively in the Mann-Whitney rank sum test.

Figure 5.3: Trust for high-income group



Note: The amount sent in period 1 for high-income subjects in high and low income inequality, according to mechanism. Error bars indicate 10% confidence intervals. Amounts are in tokens.

Figure 5.4: Expectations and trustworthiness for high-income group by Mechanism



Note: The figure shows the expectations (left column) and trustworthiness (right column) for the Random, Merit and Greed treatments respectively for high-income subjects. In the left column, the x-axis indicates the hypothetical amount sent (ranging from 0 to 300 for high-income subjects) and the y-axis indicates the expected amount returned. In the right column, the x-axis indicates the hypothetical amount received before tripling (ranging from 0 to 300) and the y-axis indicates the amount returned. Amounts are in tokens.

Table 5.7: Expectations and Trustworthiness differences between treatments for high-income group

Panel A: Expectations																		
Amount sent				20	40	60	80	100	120	140	160	180	200	220	240	260	280	300
Testing inequality	RandomHigh	-	RandomLow	-17.00**	-32.00***	-48.00***	-58.50***	-70.00***	-85.50***	-109.00***	-107.00***	-128.00***						
	MeritHigh	-	MeritLow	0.00	-12.00	-17.00	-19.00	-25.00	-23.00	-29.00	-28.00	-40.00						
	GreedHigh	-	GreedLow	5.00	2.00	-2.00	-12.50	-1.50	-5.00	-7.50	-1.00	-33.00						
Testing mechanism	RandomHigh	-	MeritHigh	-2.00	-1.00	5.00	0.50	0.00	-1.50	-2.80	1.50	-2.00	10.50	21.00	25.50	23.50	35.00	47.00
(high inequality)	RandomHigh	-	GreedHigh	-2.00	-3.00	4.00	6.00	6.50	6.50	8.20	15.50	14.00	28.50	39.50	47.50	45.50	59.00	71.00
	MeritHigh	-	GreedHigh	0.00	-2.00	-1.00	5.50	6.50	8.00	11.00	14.00	16.00	18.00	18.50	22.00	22.00	24.00	24.00
Testing mechanism	RandomLow	-	MeritLow	15.00**	19.00*	36.00***	40.00*	45.00*	61.00**	77.00**	80.00**	86.00						
(low inequality)	RandomLow	-	GreedLow	20.00***	31.00**	50.00**	52.00**	75.00***	87.00***	109.50**	121.00**	109.00**						
	MeritLow	-	GreedLow	5.00	12.00	14.00	12.00	30.00	26.00	32.50	41.00	23.00						
Panel B: Trustworthiness																		
Amount received				20	40	60	80	100	120	140	160	180	200	220	240	260	280	300
Testing inequality	RandomHigh	-	RandomLow	-17.00***	-31.00**	-41.00**	-55.00**	-75.50***	-85.00**	-92.00***	-102.00***	-111.00***						
	MeritHigh	-	MeritLow	-3.50	-5.50	-14.00	-21.00	-28.50	-32.00	-32.00	-34.00	-40.00						
	GreedHigh	-	GreedLow	6.50*	13.00	17.00	33.00**	43.50**	52.00**	59.50**	67.00**	80.50**						
Testing mechanism	RandomHigh	-	MeritHigh	-4.00	-1.00	0.00	-2.00	-6.50	-10.00	-7.00	-9.00	-12.00	-25.50	-24.00	-16.50	-16.00	-18.00	-22.50
(high inequality)	RandomHigh	-	GreedHigh	-0.50	3.00	9.00	5.00	0.00	0.00	7.50	12.00	17.50	16.50	23.50	29.50	34.50	33.00	30.00
	MeritHigh	-	GreedHigh	3.50	4.00	9.00	7.00	6.50	10.00	14.50	21.00	29.50	42.00	47.50	46.00	50.50	51.00	52.50
Testing mechanism	RandomLow	-	MeritLow	9.50	24.50**	27.00**	32.00*	40.50*	43.00	53.00*	59.00	59.00						
(low inequality)	RandomLow	-	GreedLow	23.00***	47.00***	67.00***	93.00***	119.00***	137.00***	159.00***	181.00***	209.00***						
	MeritLow	-	GreedLow	13.50***	22.50**	40.00***	61.00***	78.50***	94.00***	106.00***	122.00***	150.00***						

Note: The table reports differences between the expectations and trustworthiness (measured in tokens) in the treatment reported in the second column and the treatment reported in the third column for high-income subjects. The treatments either have the same mechanism ('Testing inequality') or the same level of inequality ('Testing mechanism'). ***, ** and * indicates significance at the 1%, 5% and 10% respectively in the Mann-Whitney rank sum test.

5.3.5 Salience and fairness

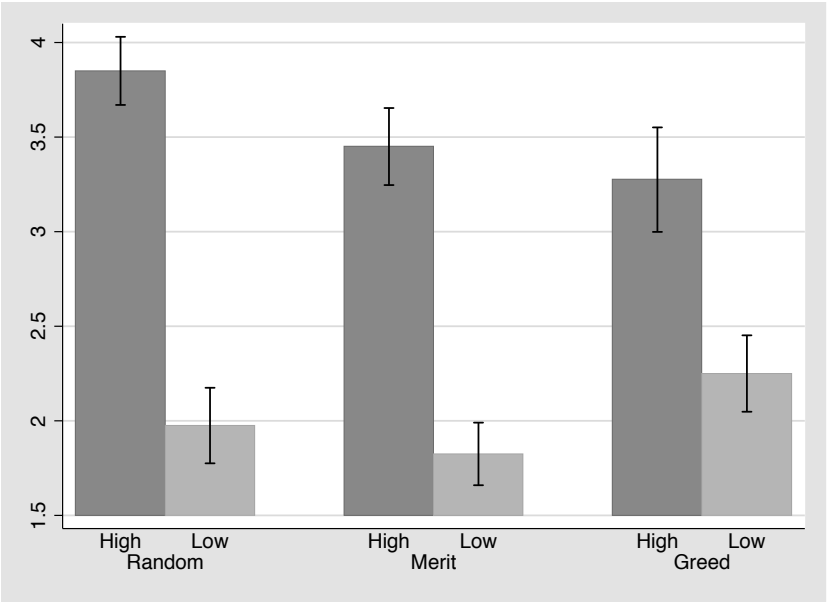
In order to determine the salience of income differences, we ask "In your opinion, were the differences between the incomes of the High Income group and the Low Income group in the Multiplier Game[10] in period 1: Very small - Small - Somewhere between small and large - Large - Very large." We translate these answers to a five-point scale (1=very small,...,5=very large) and graph the average for the six treatments in Figure 5.5. The first aspect to notice is that high inequality is always perceived as higher than low inequality. This result confirms that subjects indeed experienced, what we call, 'high inequality' as higher than 'low inequality.' Also, income differences are more salient in the Random treatment than in the Merit and Greed treatments. The difference between high and low inequality is particularly small in the Greed treatment.

To measure the perception of fairness we ask subjects the question: "Would you say that distributing the income on the basis of the chosen task from Stage 1 was: Very unfair - Unfair - Neutral - Fair - Very fair."[11] Again, we convert the answers to a five-point scale (1=very unfair,...,5=very fair). Figure 5.6 shows the perception of fairness for the low-income group. In contrast to our expectations, we see no difference in fairness perceptions in the Random treatment. Albeit insignificant, we do find differences in the Greed and Merit treatments, and in these treatments fairness is higher in low inequality than in high inequality. Next, we only look at the sample of low-income subjects that would have been in the low-income group regardless of the treatment. By focusing on this sample, we control for selection effects (see Section 5.4.1 for more details). Figure 5.7 shows the fairness for the selected sample. We now find a different pattern: there is a higher level of trust in RandomLow than in RandomHigh, while there is no difference between high and low inequality in the other treatments.

[10] The name we used in the experiment to refer to the trust game.

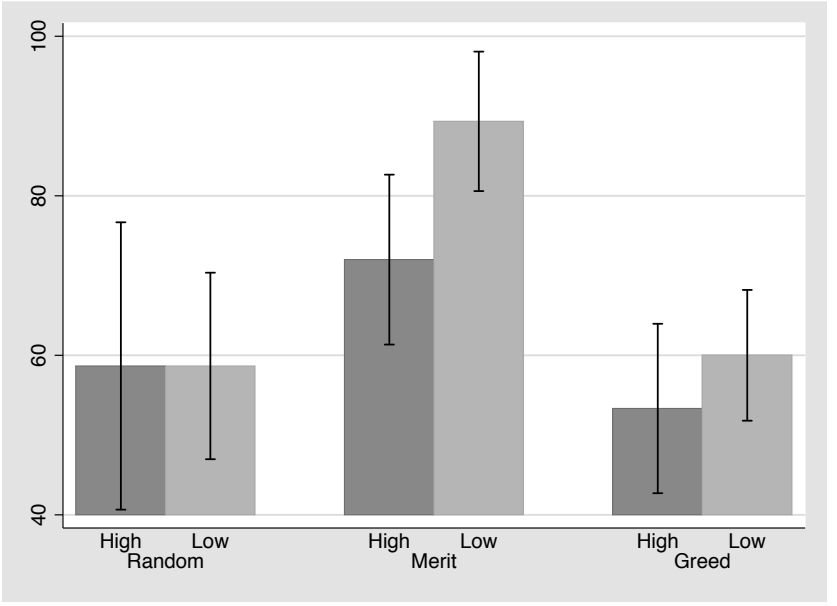
[11] This question is asked at the end of the experiment, i.e. after two periods in which subjects experienced both high and low income inequality. We still choose to split up subjects according to their experience of high and low inequality in the first period, because we saw that results for the second period were highly persistent. Nevertheless, we should be careful with the interpretation and the results should mainly be regarded as exploratory.

Figure 5.5: Perceived income inequality for low-income group



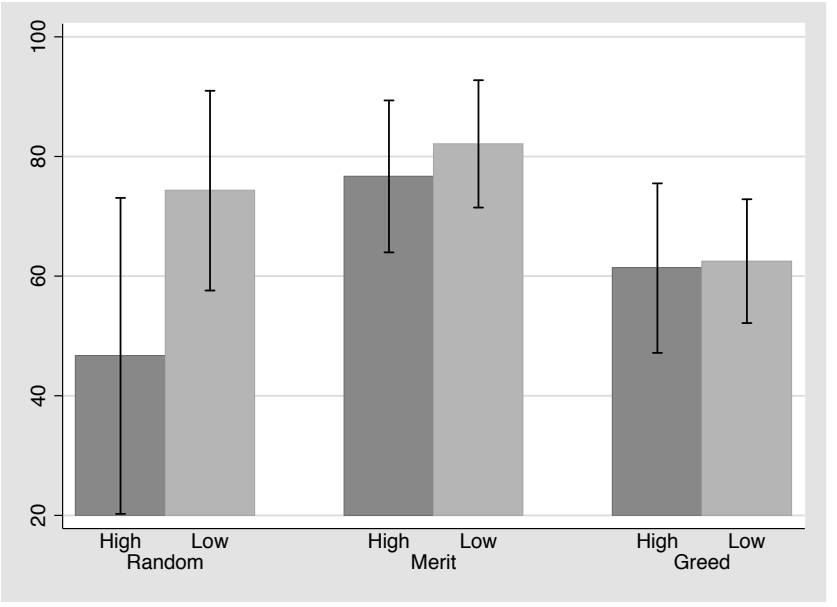
Note: This figure shows the perceived inequality in high inequality and low inequality for low-income subjects. Error bars indicate 10% confidence intervals.

Figure 5.6: Fairness for low-income group



Note: This figure shows the perceived fairness for low-income subjects. Error bars indicate 10% confidence intervals.

Figure 5.7: Fairness for low-income group, corrected for selection effects



Note: This figure shows the perceived fairness for a selected sample of low-income subjects. Error bars indicate 10% confidence intervals.

5.4 Robustness checks

5.4.1 Selection effects

Our main finding is that inequality impacts trust, but only if income is distributed randomly. Our results could be explained by the fact that, by design, the low-income subjects have different characteristics across the mechanisms. In the Merit treatment we select subjects that are the least competent in calculations or make the least effort, while in the Greed treatment we select subjects that are the least greedy. This selection process affects not only the comparison of trust levels between mechanisms, but also the analysis of inequality effects within each mechanism. After all, we might select subjects that are less responsive to differences in income inequality. We test for selection effects in two ways. We first redo our nonparametric analysis for selected samples that are comparable across mechanisms. Next, we run regressions to control for selection effects parametrically.

Subjects sometimes end up in the low-income group, while they would have been assigned to the high-income group in another treatment. We redo our nonparametric analysis for subjects that would have been in the low-income group regardless of the treatment.[12] Figure 5.8 shows that our conclusions are generally maintained if we only look at the selected sample.[13] As in our analysis of the full sample, we find a difference between high and low inequality when income is distributed randomly, but not when income is distributed based on merit or greed.[14] Our earlier conclusions generally hold for the expectations and trustworthiness of the selected sample, though not always significantly.[15]

A disadvantage of controlling for selection effects nonparametrically is that we have to drop half

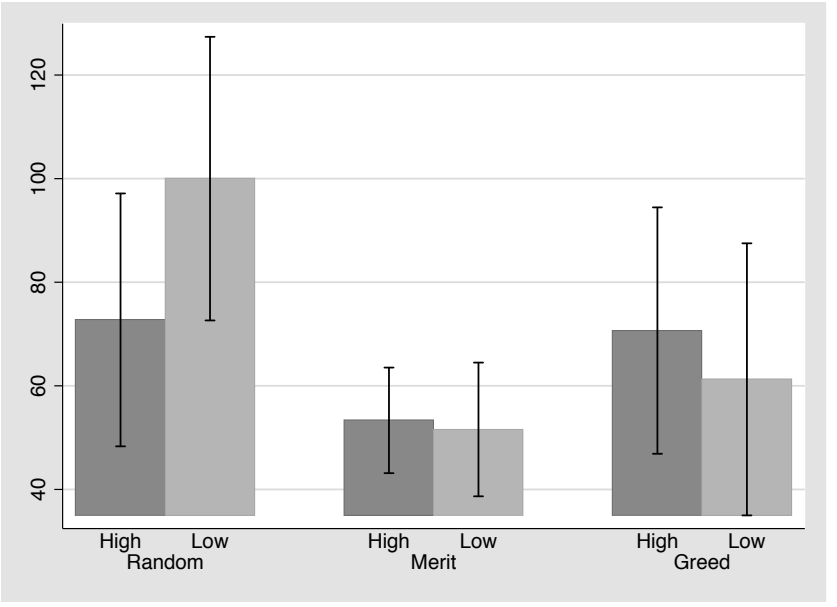
[12] First, we calculate the minimum number of correct answers in the calculation task and the minimum amount of tokens allocated to oneself in the decider task among the high-income group in the Merit and Greed treatments respectively: 12 correct answers and 500 tokens. These numbers form the 'Merit threshold' and the 'Greed threshold' to enter the high-income group. Second, we drop all low-income subjects above at least one of these thresholds. This implies that we drop 93 out of 180 observations and remain with about 50% of the original low-income sample.

[13] See Table 5.B7 for the nonparametric test results.

[14] None of the differences are statistically significant, probably due to small sample sizes.

[15] See figure 5.C5 and Table 5.B8. The only exception are the significantly higher expectations in GreedHigh than in GreedLow.

Figure 5.8: Trust for low-income group, corrected for selection effects



Note: The amount sent in period 1 for a selected sample of low-income subjects in high and low income inequality, according to mechanism. Error bars indicate 10% confidence intervals. Amounts are in tokens.

of our sample. For that reason, we also conduct a parametric analysis. Table 5.8 shows the regression results for the low-income group, controlling for performance in the calculation and decider task.[16] The coefficient of 'tokens kept in the decider task' is negative and significant at the 1% level for all dependent variables. 'Correct answers in calculation task' is insignificant in all three regressions. Figure 5.9 shows the marginal effects of the treatments according to the regression results.[17] It is clear that our main conclusion does not change: we still only find an inequality effect in the random distribution.[18] The exception is trust in the merit-based distribution, which is significantly lower in MeritLow, but this difference is not significant according to the Wald Test. Our conclusions for the high-income group also remain the same after controlling for selection parametrically.[19] To conclude, both nonparametric and parametric methods show that our results are robust against selection effects.

[16] Instead of running separate regressions for expectations and trustworthiness for each amount sent or received, we pool the observations, include dummies for amount sent/received and cluster standard errors at the individual level.

[17] The marginal effects of the treatments are the treatment effects for the average number of correct answers in the calculation (9.61) and the average number of tokens allocated to oneself in the decider task (418.15) for the low-income group.

[18] See also Wald test results at the bottom of Table 5.8)

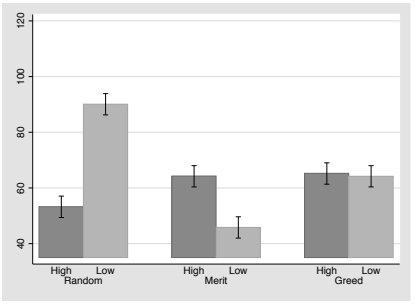
[19] See Table 5.B9 and Figure 5.C6.

Table 5.8: Regressions for low-income group

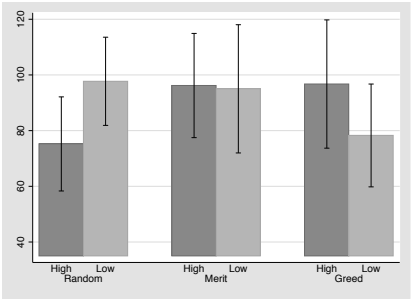
	Trust			Expectations			Trustworthiness		
	(a)	(b)	(c)	(a)	(b)	(c)	(a)	(b)	(c)
Constant	53.33***	78.90***	43.60	-1.66	121.74***	79.50**	-11.55	118.08***	35.23
	(9.35)	(23.68)	(29.16)	(10.78)	(29.22)	(33.70)	(10.71)	(32.89)	(32.49)
RandomLow	34.67***	36.83***	27.59**	26.41*	22.50	0.37	35.94**	35.79**	25.78**
	(13.23)	(13.08)	(12.37)	(15.30)	(13.99)	(10.73)	(16.74)	(15.11)	(11.78)
MeritHigh	8.67	10.96	2.03	21.31	20.97	1.22	32.27*	35.08**	24.30**
	(13.23)	(13.00)	(12.26)	(16.87)	(15.20)	(9.83)	(17.55)	(15.76)	(10.57)
MeritLow	-8.00	-7.42	-15.26	27.97	19.79	-4.16	48.47**	42.92**	32.45***
	(13.23)	(13.13)	(12.36)	(18.65)	(17.47)	(11.64)	(19.27)	(17.63)	(12.13)
GreedHigh	15.33	11.96	5.94	32.28*	21.51	13.38	28.18	15.33	3.93
	(13.23)	(13.06)	(12.21)	(19.26)	(17.33)	(14.63)	(18.02)	(15.13)	(12.23)
GreedLow	12.00	10.95	8.92	9.91	3.05	-1.84	17.44	10.72	9.34
	(13.23)	(13.01)	(12.11)	(16.08)	(15.28)	(14.94)	(18.84)	(16.82)	(15.95)
Correct answers in calculation task		2.28	2.49*		-1.69	-2.57		1.73	2.57
		(1.44)	(1.36)		(1.75)	(1.97)		(2.30)	(2.30)
Tokens kept in decider task		-0.11***	-0.04		-0.24***	-0.06		-0.34***	-0.21***
		(0.04)	(0.04)		(0.05)	(0.04)		(0.05)	(0.04)
Risk aversion			-3.00			-4.04*			2.66
			(2.19)			(2.22)			(2.16)
Trustworthiness			0.03			0.65***			
			(0.04)			(0.07)			
Expectations			0.14***						0.58***
			(0.04)						(0.07)
Obs	180	180	180	1440	1440	1440	1620	1620	1620
R-squared	0.07	0.12	0.25	0.27	0.34	0.60	0.20	0.31	0.49
Sent/received amount fixed effects	No	No	No	Yes	Yes	Yes	Yes	Yes	Yes
Clustered standard errors	No	No	No	Yes	Yes	Yes	Yes	Yes	Yes
P-values for Wald test									
Testing inequality									
MeritHigh=MeritLow	0.21	0.16	0.16	0.73	0.95	0.58	0.43	0.69	0.50
GreedHigh=GreedLow	0.80	0.94	0.81	0.24	0.31	0.38	0.60	0.80	0.75
Testing mechanism									
RandomHigh=MeritHigh	0.51	0.40	0.87	0.21	0.17	0.90	0.07	0.03	0.02
RandomHigh=GreedHigh	0.25	0.36	0.63	0.10	0.22	0.36	0.12	0.31	0.75
MeritHigh=GreedHigh	0.61	0.94	0.75	0.58	0.98	0.36	0.83	0.25	0.12
RandomLow=MeritLow	0.00	0.00	0.00	0.93	0.87	0.65	0.53	0.70	0.60
RandomLow=GreedLow	0.09	0.05	0.13	0.27	0.19	0.87	0.34	0.16	0.30
MeritLow=GreedLow	0.13	0.16	0.05	0.33	0.35	0.87	0.16	0.11	0.15

Note: OLS regressions for low-income subjects of trust, expectations and trustworthiness (measured in tokens) on treatments and control variables. ***, ** and * indicates significance at the 1%, 5% and 10% respectively. Standard errors are given in parentheses. Risk aversion is measured by the number of safe choices in the lottery from Holt and Laury (2002). The regressors expectations and trustworthiness in the regression for trust are the expectations and trustworthiness for 160 tokens sent/received.

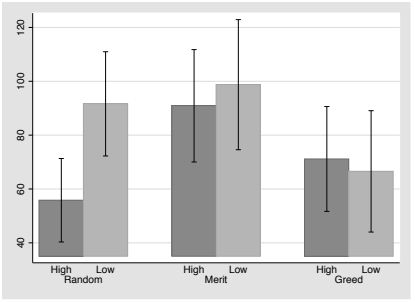
Figure 5.9: Marginal effects for low-income group, corrected for selection effects parametrically



(a) Trust



(b) Expectations



(c) Trustworthiness

5.4.2 Trust measures

What are we actually measuring with trust, expectations, and trustworthiness? In the exit survey, we ask the World Values Survey (WVS) question on trust[20]: "Generally speaking, would you say that most people can be trusted or that you cannot be too careful in dealing with people?" This question is the most commonly used survey question to measure trust. Subjects can chose one of three possible answers: "Most people can be trusted", "Cannot be too careful in dealing with people", or "Do not know." This question is commonly used in empirical studies on generalised trust. Table 5.9 shows the descriptive statistics of trust, expectations, and trustworthiness for each possible answer to the WVS trust question.[21] Our trust measure is not significantly different between those who trust and those who do not trust according to the WVS trust question. Expectations and trustworthiness, however, are significantly different between trusting and non-trusting subjects.[22] This result echoes the findings of Sapienza et al. (2013) that expectations are more correlated with WVS trust than Sender's behaviour, and that there is a correlation between trustworthiness and WVS trust. The latter can be explained by the fact that subjects use their own trustworthiness to determine their expectations of others' trustworthiness and thus their own trust.

One of the possible reasons that our measure of trust is not correlated with WVS trust is that trust is influenced by risk and other-regarding preferences. Table 5.8 shows regression results controlling for the performance in the decider task (i.e. selfishness), trustworthiness (i.e. reciprocity) and risk

[20]The full WVS questionnaire for Wave 6 can be found under the link http://www.worldvaluessurvey.org/WVSDocumentationWV6.jsp (last accessed on June 12, 2014). Question V24 reads "Generally speaking, would you say that most people can be trusted or that you need to be very careful in dealing with people?" The possible answers are "1 Most people can be trusted" and "2 Need to be very careful."

[21]Surprisingly, only 100 out of 180 (55.6%) subjects report that they believe most people can be trusted. Conditional on being Dutch, this number drops to 23.8%. In the WVS survey Wave 6, 66.1% of respondents in The Netherlands gave this answer (http://www.worldvaluessurvey.org/WVSOnline.jsp, last accessed on June 12, 2014).

[22]We choose to display expectations and trustworthiness for 160 tokens sent and received, because Sapienza et al. (2013) show that expectations and trustworthiness are correlated with WVS trust for higher amounts. In our experiment, expectations and trustworthiness are significantly different between NoTrust and Trust for all amounts.

Table 5.9: Descriptive Statistics by WVS trust category for low-income group

	NoTrust	Trust	Don'tKnow
Trust	58.20	73.00	68.50
	(4.89)	(9.10)	(8.59)
Expectation for 160 tokens sent	139.45	208.68***	174.12
	(11.85)	(19.45)	(21.37)
Trustworthiness for 160 tokens received	99.62	178.00***	146.82**
	(11.66)	(21.59)	(17.37)
Observations	100	40	40

Note: The table reports the mean of trust, expectations (for 160 tokens sent) and trustworthiness (for 160 tokens received) for the three answer options of the WVS trust question for low-income subjects. Standard errors are given in parentheses. ***, ** and * indicates a significant difference with the NoTrust category at the 1%, 5% and 10% respectively in the Mann-Whitney rank sum test.

aversion.[23] In addition, we include expectations, as the amount sent in the trust game may be explained by beliefs and preferences (Sapienza et al., 2013). Table 5.8 shows that, although the effect is slightly attenuated, trust remains significantly higher in RandomLow than in RandomHigh. Trust in the Merit and Greed treatments is still not significantly different between high and low inequality. Interestingly, we find that none of the other-regarding preferences is significant. Our conclusions are also maintained for trustworthiness, but not for expectations. The coefficient of RandomLow diminishes in the expectations regression, probably because trustworthiness is highly significant.[24] In conclusion, our findings are generally robust to inclusion of preferences and beliefs.

[23]We measure risk aversion by the number of safe choices in the lottery from Holt and Laury (2002) If we use different measures, such as the midpoint of the range of relative risk aversion for each number of safe choices, we arrive at the same qualitative conclusions.

[24]The R-squared almost doubles after including the extra regressors. This jump is completely due to inclusion of trustworthiness: If we exclude risk aversion, the R-squared is remains 0.60.

5.5 Discussion

Consistent with the empirical literature, our experimental results suggest that inequality has a negative impact on trust. However, we only find this when income is distributed randomly; inequality does not affect our outcome variables if the income distribution is based on merit or greed. In this section, we offer a potential explanation for the differential effect of distribution mechanism on the relationship: *ingroup/outgroup* effects. When income assignment is random, the only relevant facet that separates the high- and low-income group is the income differential. As the income of the high class increases, the ingroup bias relatively strengthens, and subjects want to send and return less to a member of the other group. This kind of behaviour results in lower trust and trustworthiness under higher income inequality.

Ingroup favouring is also to be expected when income is distributed based on merit or greed. However, as opposed to a society where income is distributed randomly, groups are now different in terms of income *and* characteristics (either meritorious or greedy behaviour). We posit that the class division based on characteristics is the main driver of ingroup/outgroup effects, dominating the differences in income. Several of our findings support the argument that income inequality is less important to individuals when the income distribution is nonrandom. We find that income differences are less salient in the Merit and Greed treatments than in the Random treatment. Also, fairness perceptions depend on the degree of income inequality when income is distributed randomly, but not when it is based on merit or greed. In addition, we find that trust levels are relatively low in the Merit and Greed treatments even in the case of small income differences. If ingroup/outgroup effects are indeed based solely on characteristics, it follows naturally that there is no difference in trust for different degrees of income inequality.

This explanation is supported by previous experimental studies on inequality, which have found that ingroup/outgroup effects can play a role. Lei and Vesely (2010) find evidence for ingroup favouring for rich players in a modified version of the trust game from Berg et al. (1995), but not for poor play-

ers. Their experimental design differs from ours with regards to information and mechanism: subjects know their opponent's income class and are always randomly assigned to a rich or poor position. Ku and Salmon (2013) confirm that ingroup favouring can have an impact on the tolerance for income inequality. More specifically, they show that poor subjects in their investment game are willing to transfer more to rich subjects - yielding more efficiency and higher income inequality - if they are part of the same group. Group membership is here not defined in terms of income class, but is based on other characteristics, such as performance in a test with questions from an SAT sample exam. Their design does not allow to test for the ingroup effect for characteristics that simultaneously determine group membership and income class. Nevertheless, the findings are indicative of the importance of ingroup favouring.

As Ku and Salmon (2013) is, to our knowledge, the only study to explore the role of the income distribution mechanism, it is appropriate to further compare our findings to theirs. Their experiment focuses on income inequality *as an outcome* rather than on initial (exogenous) inequality. Initial inequality is the same across treatments and it can only change as a result of behaviour in the investment game. We can therefore not compare our findings on the impact of different degrees of income inequality.

If we roughly assume that higher trust equates with higher efficiency, we *can* compare our findings with regards to the level of trust across different income distribution mechanisms to the level of investment in Ku and Salmon (2013). We focus on our results for high - rather than low - inequality since this distribution resembles more closely the income differences in Ku and Salmon (2013), where endowments for rich subjects are three times higher than for poor subjects. They find that efficiency is lower when income differences are merit- or greed-based than when there is a random assignment to income classes. In contrast, we find weak evidence that trust and trustworthiness is lower when income is distributed randomly. This result does not refute the conclusions in Ku and Salmon (2013) as our experiments differ in two important respects. First, subjects in Ku and Salmon (2013) always

play against an opponent of the other income class, meaning that ingroup bias towards members of the own income group should not affect their results. It is natural to assume that this ingroup bias is stronger in merit- and greed-based societies, since individuals have more in common with their own income class than in a society where income is assigned randomly. In case of a nonzero probability of playing against someone from the own income class, trust or efficiency might be relatively low in the randomly generated income distribution as a result of less ingroup favouring. Even if ingroup favouring does not play a role, the mere uncertainty of the status of the opponent can generate different results. Second, we are using different games to test our hypotheses. In contrast to the investment game, the trust game implies uncertainty in outcomes. Even though our belief elicitation among Senders allows us to take the uncertainty component out of the trusting behaviour, the games have different dynamics - for example the role of the second mover in the trust game - and it is therefore not obvious that subjects would behave the same.

5.6 Conclusion

This paper is a first step towards understanding how cultural and institutional factors in a society affect the relationship between income inequality and trust. Our main contribution is the investigation of the role of different income distribution mechanisms. Where present literature mainly focuses on a random distribution of income, we also analyse at merit- and greed-based distributions. When income is randomly distributed, our results suggest that individuals trust less when inequality is higher. This finding is consistent with the empirical literature. In contrast to existing studies, we also investigate cases where income is distributed on the basis of merit or greed. In such distributions a change to the level of income inequality has no significant effect on generalised trust. We show that these results hold for expected trustworthiness, which is arguably a more accurate measure of trust than Sender behaviour. Also, our findings are robust against selection effects. We hypothesise that our conclusions

can be explained by different ingroup/outgroup effects across the mechanisms.

Our findings suggest that individuals not only care about the level of inequality in their society, but also about the process behind it. This has implications for the wider debate on the effects of inequality, as well as for policies affecting redistribution. However, it is important to investigate the external validity of our research. A natural next step would be to investigate the channels that may drive our findings, particularly the influence of ingroup favouring and conflicting social norms. Such extensions could take place in the laboratory, but also in the field. In addition, theoretical models could potentially provide a deeper understanding of the channels. If our hypotheses are not rejected, future extensions could explore the effect of the income distribution mechanism on other societal phenomena linked to income inequality, such as social distance and social mobility.

5.A Instructions

WELCOME!

The experiment is run in two stages. Throughout the experiment, you will have the chance to earn tokens. **One token is equal to one euro cent.** Your earnings in Stage 1 may affect your earnings in Stage 2.

In Stage 1, you will undertake three simple, independent tasks:

- The Calculation Task
- The Decider Task
- The Roulette Task

After everyone has finished all three tasks, we will randomly select **one of the three tasks**. Only the earnings from this selected task will be paid out, and only your performance in this task will affect Stage 2. You will not be paid for the other two tasks.

At the end of Stage 2, you will be asked to fill out a short questionnaire, including one question in which you can earn extra money. After the questionnaire, you will be told your total earnings from the experiment, which consist of:

- Your earnings in Stage 1
- Your earnings in Stage 2
- Your earnings in the questionnaire

When you have finished reading, click "OK" to begin the experiment.

Task 1: CALCULATION TASK

You will be given four minutes to solve as many simple calculation problems as you can. In each problem, you will be shown two 4x4 boxes filled with a number between 1 and 100. You must find

the **largest number** in each of the two boxes, add the two numbers together, and type in your answer. Once you confirm your answer to a problem, you will be shown the next one. For each question you get right, you earn **40 tokens**.

If you want, you can use a calculator. You can access the calculator by clicking on the calculator image in the lower right corner of the screen.

Before you start, you can try one **practice round**. The result of this round does not count for your earnings.

TASK 2: DECIDER TASK

You will be anonymously paired with another participant, and you will have to split **500 tokens** between the two of you. Only **one** of you, the Decider, gets to choose how the money is split. The Decider will be randomly chosen after you both give your answers.

However, every token the Decider allocates to his or her partner is **DOUBLED**. So, if the Decider splits the money in half, he or she receives 250 tokens and the partner receives 500 tokens. The Decider's decision is final; the other person in the pair has no say in the allocation and must accept the choice.

Imagine now that you are chosen as the Decider. How much of the 500 tokens do you want to allocate to yourself and your partner? Use the slider to choose the amount you would allocate to yourself. (To use the slider use your mouse and for fine tuning use the left/right arrow keys on your keyboard.)

TASK 3: ROULETTE TASK

Our roulette wheel has slots numbered from 1 to 36 equally arranged in a circle. All you have to do is choose a number from 1 to 36. Next, a random number generator will draw the 'winning number.' Your earnings depend on how close your choice is to the winning slot, in terms of distance.

You will earn **630 minus 35 tokens for every slot your number is away from the winning number**. For example, if you choose 1 and the winning number is 36, your number is only one slot away, so you

earn 630 - 35 = 595 tokens. But if you choose 18, your number is 18 slots away and so you earn 630 - 35*18 = 0 tokens.

When you are ready, choose your number. When you know your number, you can continue to Stage 2 of the experiment.

STAGE 2

The random task chosen in Stage 1 is the [**Calculation Task / Decider Task / Roulette Task**]. Only this task from Stage 1 will be paid out. You will see your exact earnings at the end of the experiment, but on the next screen you will find out your relative performance compared to the group. This will determine your income class for Stage 2. The top 5 earners from the [**Calculation Task / Decider Task / Roulette Task**] will be in the high income class; the other 15 participants will be in the low income class.

After you find out your income class, you will play two rounds of the **"Multiplier Game"**. The Multiplier Game is played in pairs between a Sender and a Responder. You will get to play as both types. At the beginning of each round of the Multiplier Game, you and your group members will each receive your income, according to your income class, which you can use in the game. **The income received in the second round may be different from the income in the first round.**

Then, in each round, you have to make three independent decisions. For each decision you will be randomly paired with a new, anonymous partner. Thus, for every decision, **you won't know whether your partner is from the high or the low income class.**

To summarise: There are two rounds, each with three decisions, so you have six chances to earn money. On the following screen, you will learn whether you are in the high or low income group and what your income is. After that, you can start playing the Multiplier Game.

Only for high-income group

ROUND 1 ENDOWMENTS

We randomly chose a task for your group, which was the [**Calculation Task / Decider Task / Roulette Task**]. You are one of the 5 persons who [solved the **most questions correctly / allocated the most money to themselves /** who chose a number that was **closest to the winning number**] in this task. Therefore, you are assigned to the **High Income** group. In this round, Round 1, you will receive an endowment of [180 / 300]. The 15 persons who [solved the least questions correctly / allocated the least money to themselves / were assigned numbers farthest away from the winning number] are assigned to the Low Income group. They will receive an endowment of 160 tokens. See the chart below for an overview of the Round 1 endowments in your group.

Only for low-income group

ROUND 1 ENDOWMENTS

We randomly chose a task for your group, which was the [**Calculation Task / Decider Task / Roulette Task**]. You are one of the 15 persons who [solved the **least questions correctly / allocated the least money to themselves /** chose a number that was **farthest away from the winning number**] in this task. Therefore, you are assigned to the **Low Income** group. In this round, Round 1, you will receive an endowment of **160 tokens**. The 5 persons who [solved the most questions correctly / allocated the most money to themselves / were assigned numbers closest to the winning number] are assigned to the High Income group. They will receive an endowment of [180 / 300] tokens. See the chart below for an overview of the Round 1 endowments in your group.

Multiplier Game 1 - Decision 1

You are the Sender and are paired with an anonymous Responder. You can decide to send some of your income for the round to the Responder. Whatever amount you send will be **tripled** before it reaches your partner. You keep whatever amount you did not send, and then **the Responder will decide how**

much of what they received to return to you. The Responder keeps whatever is not returned.
Use the slider to indicate the amount you wish to send to the Receiver. You can only send multiples of 20 tokens. (Use your mouse or the left/right arrow keys on your keyboard.)

Multiplier Game 1 - Decision 2

In Decision 2, we want you to guess the behaviour of the Responder. Each responder has to decide how much to return to the Sender *for each possible amount* they could have received - that is, for each multiple of 20 tokens that could be chosen.
Now, imagine you had sent different amounts. How much do you think the Responder would return to you?
Indicate for *each possible amount that could have been sent* how much you guess the Responder would send back. Your earnings will be based on how closely your estimates match the Responder's behaviour. We will choose one of the choices at random, and if your estimate matches with the Responder's chosen amount to return (with a 10% margin of error), you will earn **100 tokens**.
(Use the 'TAB' key to quickly move your cursor to the next box.)

Multiplier Game 1 - Decision 3

In the third decision, you will be repaired with a different, anonymous partner, but this time they will be the Sender and you will be the Responder. The Sender will decide how much of his/her income to send to you, which will be multiplied by three. You must decide how much of this amount to send back to them, and you will earn whatever is remaining.
Indicate how much you wish to send back *for each possible amount* you might receive from the Sender. We will compare your choices with how much the Sender decided to send, and only your corresponding choice to that amount will be played out.
Remember, **you don't know whether you are paired with someone from the high or the low income**

class. However, for Sender amounts *above* the maximum low income amount, you can deduce that only someone from the high income class could send them.

Only for high-income group

ROUND 2 ENDOWMENTS

PLEASE READ CAREFULLY.

This is Round 2. Again, you are assigned income according to your income group. Remember that you are in the **High Income** group because you are one of the 5 persons who [solved the most questions correctly in the **Calculation Task /** allocated the most money to themselves in the **Decider Task /** chose a number that was closest to the winning number in the **Roulette Task**]. In the previous round you received an endowment of [180 / 300] tokens. In this round, Round 2, you will receive an endowment of [**300 tokens / 180 tokens**]. The 15 persons in the Low Income group will receive an endowment of 160 tokens, the same amount as they received in the previous round. See the chart below for an overview of the Round 2 endowments in your group.

Only for low-income group

ROUND 2 ENDOWMENTS

PLEASE READ CAREFULLY.

This is Round 2. Again, you are assigned income according to your income group. Remember that you are in the **Low Income** group because you are one of the 15 persons who [solved the least questions correctly in the **Calculation Task /** allocated the least money to themselves in the **Decider Task /** chose a number that was farthest away from the winning number in the **Roulette Task**. In the previous round you received an endowment of 160 tokens. In this round, Round 2, you will again receive an endowment of **160 tokens**. The 5 persons in the High Income group will receive an endowment of [300 / 180] tokens, while they received [180 / 300] tokens in the previous round. See the chart below for an

overview of the Round 2 endowments in your group.

Multiplier Game 2 - Decision 1

You are the Sender. Use the slider to indicate the amount you wish to send to the Receiver. You can only send multiples of 20 tokens.

NOTE: To use the slider use your mouse and for fine tuning use the left/right arrow keys on your keyboard.

Multiplier Game 2 - Decision 2

Again, you are the Sender. Indicate for *each possible amount* how much you expect to get back from the Receiver.

Multiplier Game 2 - Decision 3

Now, you are the Receiver. Indicate for *each possible amount* how much you would send back to the Receiver.

You have completed both stages of the experiment. You now have the opportunity to increase your earnings by filling out some preferences. Your earnings in this part of the experiment depend only on your own decisions and they will be added to your previous earnings and paid to you in cash at the end of the experiment.

5.B Tables

Table 5.B1: Descriptive Statistics for low-income group in period 1

	All	RandomHigh	RandomLow	MeritHigh	MeritLow	GreedHigh	GreedLow
Trust	63.78	53.33	88.00	62.00	45.33	68.67	65.33
	(52.25)	(41.80)	(56.96)	(46.79)	(43.61)	(59.58)	(55.82)
Expectation for 20 tokens sent	17.99	14.40	18.00	19.03	18.00	20.83	17.67
	(17.84)	(17.09)	(14.24)	(17.38)	(21.24)	(19.12)	(18.13)
Expectation for 40 tokens sent	38.76	30.57	40.83	36.83	41.47	46.17	36.67
	(33.86)	(33.89)	(26.78)	(31.58)	(39.55)	(38.45)	(31.98)
Expectation for 60 tokens sent	60.91	47.23	64.17	61.20	63.67	69.17	60.00
	(50.59)	(45.09)	(38.06)	(48.12)	(59.04)	(57.21)	(54.39)
Expectation for 80 tokens sent	80.63	66.23	87.50	77.70	88.00	93.33	71.00
	(64.22)	(58.12)	(49.07)	(62.53)	(76.27)	(76.80)	(58.80)
Expectation for 100 tokens sent	100.24	78.23	112.50	99.70	107.33	115.83	87.83
	(78.13)	(73.74)	(60.70)	(79.72)	(90.05)	(94.47)	(63.59)
Expectation for 120 tokens sent	119.38	92.90	130.50	122.87	131.00	134.33	104.67
	(92.60)	(87.52)	(71.95)	(92.93)	(105.16)	(113.16)	(78.29)
Expectation for 140 tokens sent	138.45	108.57	149.17	142.00	154.00	158.47	118.50
	(110.60)	(104.24)	(91.12)	(108.35)	(123.95)	(133.26)	(96.63)
Expectation for 160 tokens sent	162.54	123.57	170.33	172.83	182.00	181.83	144.67
	(125.90)	(120.02)	(105.09)	(119.41)	(138.30)	(153.02)	(112.88)
Trustworthiness for 20 tokens received	15.50	12.67	18.50	14.50	19.33	15.67	12.33
	(17.03)	(16.39)	(17.18)	(14.40)	(19.11)	(16.54)	(18.32)
Trustworthiness for 40 tokens received	30.75	25.00	33.50	30.67	40.67	30.67	24.00
	(31.73)	(27.51)	(27.45)	(29.35)	(36.85)	(30.95)	(36.54)
Trustworthiness for 60 tokens received	47.33	35.83	50.83	49.33	63.67	47.67	36.67
	(48.40)	(39.79)	(44.18)	(45.40)	(54.93)	(48.54)	(54.22)
Trustworthiness for 80 tokens received	64.25	47.00	70.17	64.33	83.00	66.33	54.67
	(63.13)	(52.40)	(52.86)	(59.05)	(72.55)	(62.94)	(74.45)
Trustworthiness for 100 tokens received	80.39	53.33	90.17	81.33	103.33	83.33	70.83
	(79.45)	(66.97)	(71.59)	(75.83)	(90.07)	(79.80)	(87.28)
Trustworthiness for 120 tokens received	95.72	61.00	106.50	104.00	124.33	95.00	83.50
	(94.47)	(76.53)	(85.29)	(90.61)	(107.40)	(95.94)	(102.80)
Trustworthiness for 140 tokens received	112.79	70.73	125.17	125.33	142.33	113.00	100.17
	(110.38)	(91.07)	(98.66)	(107.73)	(126.67)	(114.02)	(115.34)
Trustworthiness for 160 tokens received	127.53	84.40	138.67	142.67	152.60	134.33	112.50
	(123.67)	(103.70)	(115.01)	(119.16)	(140.91)	(127.00)	(130.44)
Trustworthiness for 180 tokens received	137.91	78.77	158.67	147.00	175.67	136.33	131.00
	(139.41)	(110.99)	(122.53)	(142.49)	(157.67)	(138.30)	(150.55)
Trustworthiness for 200 tokens received	68.27	91.47		163.33		154.83	
	(125.20)	(126.81)		(157.84)		(153.73)	
Trustworthiness for 220 tokens received	73.53	98.17		178.33		164.67	
	(135.13)	(138.49)		(171.33)		(163.74)	
Trustworthiness for 240 tokens received	79.98	104.53		194.67		180.67	
	(147.51)	(151.14)		(187.50)		(178.15)	
Trustworthiness for 260 tokens received	87.27	119.27		212.33		192.00	
	(158.78)	(163.11)		(202.05)		(189.62)	
Trustworthiness for 280 tokens received	94.72	133.67		227.00		207.67	
	(171.45)	(179.26)		(215.87)		(204.92)	
Trustworthiness for 300 tokens received	105.89	146.00		253.00		236.33	
	(186.67)	(194.91)		(229.02)		(219.52)	
Observations	180	30	30	30	30	30	30

Note: The table reports the mean of the main outcome variables for low-income subjects in period 1. Amounts are in tokens. Standard deviations are given in parentheses.

Table 5.B2: Descriptive Statistics for low-income group in period 2

	All	RandomHigh	RandomLow	MeritHigh	MeritLow	GreedHigh	GreedLow
Trust	62.00	48.67	76.00	63.33	49.33	69.33	65.33
	(50.79)	(49.74)	(49.10)	(50.40)	(45.71)	(58.19)	(48.97)
Expectation for 20 tokens sent	16.99	12.90	19.33	15.53	20.50	17.50	16.17
	(17.97)	(17.89)	(16.60)	(15.52)	(20.61)	(19.24)	(17.99)
Expectation for 40 tokens sent	34.85	29.07	39.00	33.03	41.83	32.33	33.83
	(34.99)	(31.63)	(31.22)	(31.93)	(42.11)	(37.20)	(35.95)
Expectation for 60 tokens sent	53.43	47.23	58.33	52.37	62.67	50.50	49.50
	(51.17)	(44.63)	(44.42)	(46.92)	(62.80)	(55.28)	(53.05)
Expectation for 80 tokens sent	73.24	61.73	82.50	66.03	92.17	68.67	68.33
	(69.12)	(59.58)	(58.50)	(63.89)	(87.08)	(72.76)	(69.73)
Expectation for 100 tokens sent	90.68	74.57	98.67	84.70	110.33	89.33	86.50
	(83.76)	(75.16)	(71.34)	(80.17)	(100.57)	(90.74)	(83.70)
Expectation for 120 tokens sent	107.88	86.90	119.33	103.20	130.00	103.33	104.50
	(98.62)	(89.31)	(83.86)	(96.62)	(120.52)	(108.13)	(91.18)
Expectation for 140 tokens sent	124.91	97.57	136.67	124.53	151.33	121.17	118.17
	(115.38)	(104.20)	(99.94)	(112.49)	(139.28)	(127.10)	(106.82)
Expectation for 160 tokens sent	143.54	113.40	148.50	143.87	175.33	140.83	139.33
	(129.72)	(118.82)	(104.05)	(129.11)	(158.41)	(143.43)	(120.61)
Trustworthiness for 20 tokens received	15.49	11.67	17.30	15.17	20.83	17.00	11.00
	(17.44)	(17.58)	(16.34)	(15.11)	(20.00)	(17.79)	(16.89)
Trustworthiness for 40 tokens received	30.41	22.83	33.13	31.00	41.00	32.50	22.00
	(33.12)	(29.99)	(30.76)	(30.55)	(38.81)	(33.65)	(33.05)
Trustworthiness for 60 tokens received	47.66	37.67	50.97	48.67	64.33	49.50	34.83
	(49.59)	(44.62)	(45.14)	(45.47)	(61.51)	(49.42)	(47.75)
Trustworthiness for 80 tokens received	62.44	48.50	67.63	63.67	80.00	67.33	47.50
	(63.79)	(58.91)	(60.44)	(59.10)	(73.30)	(66.02)	(62.93)
Trustworthiness for 100 tokens received	78.55	57.67	87.47	82.33	100.00	83.00	60.83
	(78.96)	(73.61)	(73.44)	(76.04)	(90.48)	(82.51)	(74.35)
Trustworthiness for 120 tokens received	94.05	69.50	105.13	100.33	118.33	99.00	72.00
	(93.58)	(88.87)	(88.41)	(89.92)	(108.19)	(97.67)	(84.18)
Trustworthiness for 140 tokens received	109.72	78.50	119.13	119.67	139.67	118.00	83.33
	(109.33)	(104.40)	(103.99)	(103.54)	(125.27)	(115.68)	(96.72)
Trustworthiness for 160 tokens received	127.58	94.17	139.80	136.33	156.37	143.00	95.83
	(122.60)	(117.33)	(114.40)	(117.58)	(143.28)	(130.28)	(105.14)
Trustworthiness for 180 tokens received	141.19	104.00	154.47	156.00	153.37	174.17	105.17
	(136.86)	(133.59)	(129.52)	(129.44)	(147.45)	(162.35)	(108.04)
Trustworthiness for 200 tokens received	77.39		170.80		173.37		120.17
	(127.10)		(144.22)		(161.24)		(117.84)
Trustworthiness for 220 tokens received	81.72		177.30		185.70		127.33
	(136.16)		(159.97)		(174.46)		(122.08)
Trustworthiness for 240 tokens received	91.19		203.63		206.37		137.17
	(154.37)		(198.61)		(190.21)		(129.08)
Trustworthiness for 260 tokens received	97.89		223.13		221.03		143.17
	(165.85)		(213.40)		(205.91)		(134.44)
Trustworthiness for 280 tokens received	102.89		227.13		233.70		156.50
	(169.11)		(196.93)		(215.07)		(148.25)
Trustworthiness for 300 tokens received	113.39		251.30		253.70		175.33
	(182.42)		(204.22)		(231.61)		(161.71)
Observations	180	30	30	30	30	30	30

Note: The table reports the mean of the main outcome variables for low-income subjects in period 2. Amounts are in tokens. Standard deviations are given in parentheses.

Table 5.B3: Descriptive Statistics for high-income group in period 1

	All	RandomHigh	RandomLow	MeritHigh	MeritLow	GreedHigh	GreedLow
Trust	78.33	140.00	62.00	84.00	48.00	78.00	58.00
	(81.45)	(101.11)	(46.62)	(96.06)	(55.14)	(92.59)	(64.94)
Expectation for 20 tokens sent	11.33	8.00	25.00	10.00	10.00	10.00	5.00
	(13.71)	(10.33)	(15.09)	(14.14)	(11.55)	(13.33)	(10.80)
Expectation for 40 tokens sent	26.33	18.00	50.00	19.00	31.00	21.00	19.00
	(26.87)	(19.89)	(25.39)	(28.46)	(19.69)	(28.46)	(28.46)
Expectation for 60 tokens sent	43.17	35.00	83.00	30.00	47.00	31.00	33.00
	(39.42)	(28.77)	(33.02)	(43.46)	(31.99)	(41.75)	(35.92)
Expectation for 80 tokens sent	61.33	48.50	107.00	48.00	67.00	42.50	55.00
	(51.39)	(36.06)	(40.57)	(54.53)	(44.98)	(51.60)	(59.49)
Expectation for 100 tokens sent	77.92	64.00	134.00	64.00	89.00	57.50	59.00
	(60.23)	(39.78)	(53.17)	(65.69)	(54.46)	(61.88)	(57.82)
Expectation for 120 tokens sent	92.75	75.50	161.00	77.00	100.00	69.00	74.00
	(71.83)	(46.69)	(62.26)	(79.31)	(67.33)	(69.35)	(72.60)
Expectation for 140 tokens sent	111.62	89.20	198.00	92.00	121.00	81.00	88.50
	(89.66)	(58.40)	(66.80)	(92.59)	(83.99)	(82.25)	(106.88)
Expectation for 160 tokens sent	128.42	111.50	218.00	110.00	138.00	96.00	97.00
	(100.51)	(76.31)	(78.57)	(107.70)	(94.14)	(94.54)	(113.34)
Expectation for 180 tokens sent	153.50	124.00	252.00	126.00	166.00	110.00	143.00
	(112.88)	(63.63)	(90.65)	(122.85)	(106.69)	(101.54)	(139.61)
Expectation for 200 tokens sent	129.55	155.50		145.00		127.00	
	(110.68)	(72.89)		(134.10)		(116.81)	
Expectation for 220 tokens sent	148.94	184.00		163.00		144.50	
	(126.43)	(89.22)		(149.67)		(132.17)	
Expectation for 240 tokens sent	163.79	204.50		179.00		157.00	
	(139.12)	(94.65)		(164.01)		(147.88)	
Expectation for 260 tokens sent	179.55	220.50		197.00		175.00	
	(148.56)	(88.83)		(180.25)		(158.48)	
Expectation for 280 tokens sent	196.97	248.00		213.00		189.00	
	(161.22)	(97.27)		(194.82)		(169.74)	
Expectation for 300 tokens sent	213.33	274.00		227.00		203.00	
	(173.38)	(110.07)		(209.45)		(176.95)	
Trustworthiness for 20 tokens received	9.83	6.00	23.00	10.00	13.50	6.50	0.00
	(12.65)	(9.66)	(10.59)	(14.14)	(12.92)	(11.56)	(0.00)
Trustworthiness for 40 tokens received	21.25	18.00	49.00	19.00	24.50	15.00	2.00
	(26.11)	(22.01)	(24.24)	(28.46)	(24.55)	(24.61)	(6.32)
Trustworthiness for 60 tokens received	33.33	30.00	71.00	30.00	44.00	21.00	4.00
	(37.68)	(29.06)	(34.46)	(43.46)	(33.40)	(35.50)	(12.65)
Trustworthiness for 80 tokens received	48.17	42.00	97.00	44.00	65.00	37.00	4.00
	(49.28)	(37.06)	(45.72)	(56.61)	(43.53)	(43.22)	(12.65)
Trustworthiness for 100 tokens received	61.75	49.50	125.00	56.00	84.50	49.50	6.00
	(62.24)	(51.23)	(50.83)	(71.83)	(50.80)	(56.98)	(18.97)
Trustworthiness for 120 tokens received	76.17	62.00	147.00	72.00	104.00	62.00	10.00
	(74.22)	(59.59)	(63.95)	(83.64)	(65.35)	(66.63)	(31.62)
Trustworthiness for 140 tokens received	89.58	79.00	171.00	86.00	118.00	71.50	12.00
	(86.37)	(71.56)	(73.70)	(97.66)	(76.85)	(77.53)	(37.95)
Trustworthiness for 160 tokens received	102.50	92.00	194.00	101.00	135.00	80.00	13.00
	(97.50)	(78.85)	(79.47)	(113.08)	(90.34)	(89.32)	(28.30)
Trustworthiness for 180 tokens received	116.92	107.00	218.00	119.00	159.00	89.50	9.00
	(110.48)	(85.64)	(84.43)	(130.17)	(104.29)	(100.79)	(19.12)
Trustworthiness for 200 tokens received	58.75	114.50		140.00		98.00	
	(101.43)	(98.11)		(144.76)		(113.41)	
Trustworthiness for 220 tokens received	65.58	131.00		155.00		107.50	
	(111.61)	(101.92)		(160.92)		(124.30)	
Trustworthiness for 240 tokens received	73.08	150.50		167.00		121.00	
	(122.40)	(110.92)		(176.70)		(133.70)	
Trustworthiness for 260 tokens received	79.92	166.00		182.00		131.50	
	(133.55)	(120.76)		(193.15)		(144.84)	
Trustworthiness for 280 tokens received	87.50	180.00		198.00		147.00	
	(145.29)	(132.66)		(210.23)		(155.14)	
Trustworthiness for 300 tokens received	93.00	188.50		211.00		158.50	
	(157.07)	(156.67)		(225.90)		(164.15)	
Observations	60	10	10	10	10	10	10

Note: The table reports the mean of the main outcome variables for high-income subjects in period 1. Amounts are in tokens. Standard deviations are given in parentheses.

Table 5.B4: Descriptive Statistics for high-income group in period 2

	All	RandomHigh	RandomLow	MeritHigh	MeritLow	GreedHigh	GreedLow
Trust	58.00	50.00	94.00	50.00	58.00	38.00	58.00
	(69.15)	(58.31)	(80.03)	(58.31)	(63.56)	(53.71)	(95.89)
Expectation for 20 tokens sent	8.67	6.00	17.50	11.00	8.00	7.50	2.00
	(12.03)	(9.66)	(14.77)	(13.70)	(10.33)	(12.30)	(6.32)
Expectation for 40 tokens sent	21.08	16.00	48.50	20.00	21.00	15.00	6.00
	(25.13)	(18.38)	(25.39)	(27.89)	(20.79)	(24.61)	(13.50)
Expectation for 60 tokens sent	31.67	26.00	67.50	31.00	33.50	22.50	9.50
	(36.90)	(26.75)	(37.51)	(42.80)	(30.56)	(36.91)	(23.86)
Expectation for 80 tokens sent	48.83	37.00	95.50	53.00	52.50	42.00	13.00
	(49.23)	(34.01)	(48.10)	(53.14)	(48.49)	(47.56)	(29.83)
Expectation for 100 tokens sent	64.17	50.00	118.50	74.00	72.00	51.50	19.00
	(59.39)	(42.16)	(56.86)	(68.18)	(52.87)	(57.74)	(34.79)
Expectation for 120 tokens sent	77.67	57.50	146.50	93.00	81.00	60.00	28.00
	(71.37)	(52.35)	(66.92)	(82.47)	(59.90)	(70.71)	(41.31)
Expectation for 140 tokens sent	94.25	66.00	167.00	114.00	94.00	70.50	54.00
	(83.16)	(61.68)	(74.84)	(96.29)	(69.95)	(80.71)	(75.01)
Expectation for 160 tokens sent	111.17	82.50	193.50	134.00	114.00	83.00	60.00
	(94.95)	(70.84)	(85.31)	(111.38)	(85.27)	(91.90)	(76.59)
Expectation for 180 tokens sent	128.08	96.00	221.00	147.00	140.00	93.50	71.00
	(106.56)	(76.04)	(93.74)	(125.35)	(101.43)	(102.47)	(84.39)
Expectation for 200 tokens sent	159.19		250.50		160.00		83.00
	(128.69)		(105.32)		(115.57)		(110.56)
Expectation for 220 tokens sent	175.00		271.00		173.00		98.50
	(137.37)		(114.35)		(124.37)		(119.21)
Expectation for 240 tokens sent	195.65		297.00		197.50		112.00
	(156.25)		(126.50)		(144.71)		(144.90)
Expectation for 260 tokens sent	208.71		325.00		207.00		115.00
	(168.20)		(138.34)		(150.71)		(151.09)
Expectation for 280 tokens sent	227.10		349.00		231.50		123.50
	(181.88)		(148.88)		(170.98)		(157.09)
Expectation for 300 tokens sent	246.77		375.00		250.00		140.00
	(195.57)		(160.78)		(184.51)		(172.88)
Trustworthiness for 20 tokens received	9.58	7.00	23.00	10.00	10.00	7.50	0.00
	(13.41)	(9.49)	(15.67)	(14.14)	(14.14)	(11.37)	(0.00)
Trustworthiness for 40 tokens received	20.67	15.00	46.00	19.00	27.00	15.00	2.00
	(25.37)	(20.68)	(25.03)	(28.46)	(24.06)	(22.73)	(6.32)
Trustworthiness for 60 tokens received	32.83	28.00	66.00	30.00	41.00	26.00	6.00
	(37.11)	(29.74)	(35.96)	(43.46)	(36.04)	(33.48)	(18.97)
Trustworthiness for 80 tokens received	45.17	32.00	94.00	44.00	57.00	40.00	4.00
	(50.40)	(39.38)	(48.12)	(57.00)	(49.90)	(44.47)	(12.65)
Trustworthiness for 100 tokens received	59.67	39.00	117.00	66.00	71.50	51.50	13.00
	(62.48)	(47.71)	(59.45)	(71.06)	(62.01)	(55.18)	(31.99)
Trustworthiness for 120 tokens received	70.42	48.00	139.50	82.00	83.00	61.00	9.00
	(75.02)	(56.13)	(72.13)	(89.29)	(72.43)	(66.57)	(20.25)
Trustworthiness for 140 tokens received	84.25	55.00	163.50	98.00	99.50	75.50	14.00
	(87.66)	(67.21)	(83.27)	(102.18)	(87.19)	(76.47)	(32.73)
Trustworthiness for 160 tokens received	98.00	66.00	183.50	113.00	124.50	89.00	12.00
	(99.11)	(77.06)	(94.05)	(118.51)	(93.35)	(86.72)	(27.00)
Trustworthiness for 180 tokens received	112.17	77.50	208.00	125.00	140.00	103.50	19.00
	(110.64)	(87.35)	(103.37)	(133.94)	(101.00)	(97.07)	(45.33)
Trustworthiness for 200 tokens received	69.17		236.50		163.50		15.00
	(117.57)		(120.32)		(121.61)		(33.75)
Trustworthiness for 220 tokens received	77.00		261.00		179.00		22.00
	(129.83)		(135.60)		(130.76)		(52.03)
Trustworthiness for 240 tokens received	86.33		283.50		207.50		27.00
	(143.00)		(146.89)		(141.56)		(64.64)
Trustworthiness for 260 tokens received	91.50		306.00		218.00		25.00
	(152.11)		(157.92)		(146.88)		(58.74)
Trustworthiness for 280 tokens received	100.42		334.50		246.00		22.00
	(168.26)		(174.46)		(167.15)		(48.49)
Trustworthiness for 300 tokens received	99.50		359.00		218.00		20.00
	(174.81)		(183.63)		(185.40)		(42.16)
Observations	60	10	10	10	10	10	10

Note: The table reports the mean of the main outcome variables for high-income subjects in period 2. Amounts are in tokens. Standard deviations are given in parentheses.

Table 5.B5: Trust differences between treatments for low-income group in second period

				Difference
Testing inequality	RandomHigh	-	RandomLow	-27.33
	MeritHigh	-	MeritLow	14.00
	GreedHigh	-	GreedLow	4.00
Testing mechanism	RandomHigh	-	MeritHigh	-14.67
(high inequality)	RandomHigh	-	GreedHigh	-20.67
	MeritHigh	-	GreedHigh	-6.00
Testing mechanism	RandomLow	-	MeritLow	26.67
(low inequality)	RandomLow	-	GreedLow	10.67
	MeritLow	-	GreedLow	-16.00

Note: The table reports differences in trust (measured in tokens) in period 2 between the treatment reported in the second column and the treatment reported in the third column for low-income subjects. The treatments either have the same mechanism ('Testing inequality') or the same level of inequality ('Testing mechanism'). The variable Trust is the amount sent in the Trust Game. ***, ** and * indicates significance at the 1%, 5% and 10% respectively in the Mann-Whitney rank sum test.

Table 5.B6: Expectations and Trustworthiness differences between treatments for low-income group in second period

Panel A: Expectations

Amount sent				20	40	60	80	100	120	140	160
Testing inequality	RandomHigh	-	RandomLow	-6.43*	-9.93	-11.10	-20.77	-24.10	-32.43*	-39.10*	-35.10
	MeritHigh	-	MeritLow	-4.97	-8.80	-10.30	-26.13	-25.63	-26.80	-26.80	-31.47
	GreedHigh	-	GreedLow	1.33	-1.50	1.00	0.33	2.83	-1.17	3.00	1.50
Testing mechanism	RandomHigh	-	MeritHigh	-2.63	-3.97	-5.13	-4.30	-10.13	-16.30	-26.97	-30.47
(high inequality)	RandomHigh	-	GreedHigh	-4.60	-3.27	-3.27	-6.93	-14.77	-16.43	-23.60	-27.43
	MeritHigh	-	GreedHigh	-1.97	0.70	1.87	-2.63	-4.63	-0.13	3.37	3.03
Testing mechanism	RandomLow	-	MeritLow	-1.17	-2.83	-4.33	-9.67	-11.67	-10.67	-14.67	-26.83
(low inequality)	RandomLow	-	GreedLow	3.17	5.17	8.83	14.17	12.17	14.83	18.50	9.17
	MeritLow	-	GreedLow	4.33	8.00	13.17	23.83	23.83	25.50	33.17	36.00

Panel B: Trustworthiness

Amount received				20	40	60	80	100	120	140	160	180	200	220	240	260	280	300
Testing inequality	RandomHigh	-	RandomLow	-5.63*	-10.30	-13.30	-19.13	-29.80*	-35.63*	-40.63*	-45.63*	-50.47*						
	MeritHigh	-	MeritLow	-5.67	-10.00	-15.67	-16.33	-17.67	-18.00	-20.00	-20.03	2.63						
	GreedHigh	-	GreedLow	6.00	10.50	14.67	19.83	22.17	27.00	34.67	47.17	69.00*						
Testing mechanism	RandomHigh	-	MeritHigh	-3.50	-8.17	-11.00	-15.17	-24.67	-30.83	-41.17*	-42.17	-52.00*						
(high inequality)	RandomHigh	-	GreedHigh	-5.33	-9.67	-11.83	-18.83	-25.33	-29.50	-39.50	-48.83*	-70.17*						
	MeritHigh	-	GreedHigh	-1.83	-1.50	-0.83	-3.67	-0.67	1.33	1.67	-6.67	-18.17						
Testing mechanism	RandomLow	-	MeritLow	-3.53	-7.87	-13.37	-12.37	-12.53	-13.20	-20.53	-16.57	1.10	-2.57	-8.40	-2.73	2.10	-6.57	-2.40
(low inequality)	RandomLow	-	GreedLow	6.30*	11.13*	16.13*	20.13*	26.63*	33.13*	35.80*	43.97*	49.30*	50.63*	49.97	66.47	79.97	70.63	75.97
	MeritLow	-	GreedLow	9.83**	19.00**	29.50**	32.50**	39.17**	46.33*	56.33*	60.53*	48.20	53.20	58.37	69.20	77.87	77.20	78.37

Note: The table reports differences between the expectations and trustworthiness (measured in tokens) in the treatment reported in the second column and the treatment reported in the third column for low-income subjects in the second period. The treatments either have the same mechanism ('Testing inequality') or the same level of inequality ('Testing mechanism'). ***, ** and * indicates significance at the 1%, 5% and 10% respectively in the Mann-Whitney rank sum test.

Table 5.B7: Trust differences between treatments for low-income group, below Merit and Greed threshold

				Difference
Testing inequality	RandomHigh	-	RandomLow	-27.27
	MeritHigh	-	MeritLow	1.75
	GreedHigh	-	GreedLow	9.42
Testing mechanism	RandomHigh	-	MeritHigh	19.39
(high inequality)	RandomHigh	-	GreedHigh	2.06
	MeritHigh	-	GreedHigh	-17.33
Testing mechanism	RandomLow	-	MeritLow	48.42*
(low inequality)	RandomLow	-	GreedLow	38.75*
	MeritLow	-	GreedLow	-9.67

Note: The table reports differences in trust (measured in tokens) between the treatment reported in the second column and the treatment reported in the third column for low-income subjects below the Merit and Greed threshold. The treatments either have the same mechanism ('Testing inequality') or the same level of inequality ('Testing mechanism'). The variable Trust is the amount sent in the trust game. ***, ** and * indicates significance at the 1%, 5% and 10% respectively in the Mann-Whitney rank sum test.

Table 5.B8: Expectations and Trustworthiness differences between treatments for low-income group, below Merit and Greed threshold

Panel A: Expectations

Amount sent				20	40	60	80	100	120	140	160
Testing inequality	RandomHigh	-	RandomLow	-7.60	-12.21	-13.38	-12.21	-18.90	-22.86	-19.55	-16.40
	MeritHigh	-	MeritLow	6.37	0.18	3.33	-2.28	8.55	11.89	3.07	8.55
	GreedHigh	-	GreedLow	8.67	18.04	25.04	49.79	68.31*	74.04*	96.42*	94.50
Testing mechanism	RandomHigh	-	MeritHigh	-14.45**	-16.14	-22.42	-20.30	-35.68	-50.83*	-42.88	-55.68
(high inequality)	RandomHigh	-	GreedHigh	-14.12*	-22.30	-29.76	-40.30	-56.18	-64.67	-77.21	-80.18
	MeritHigh	-	GreedHigh	0.33	-6.17	-7.33	-20.00	-20.50	-13.83	-34.33	-24.50
Testing mechanism	RandomLow	-	MeritLow	-0.49	-3.74	-5.71	-10.38	-8.23	-16.09	-20.26	-30.73
(low inequality)	RandomLow	-	GreedLow	2.14	7.95	8.66	21.70	31.03	32.23	38.75	30.71
	MeritLow	-	GreedLow	2.63	11.69	14.38	32.07	39.26	48.32*	59.01	61.45*

Panel B: Trustworthiness

Amount received				20	40	60	80	100	120	140	160	180	200	220	240	260	280	300
Testing inequality	RandomHigh	-	RandomLow	-5.97	-11.10	-18.77	-25.39	-35.65	-50.71	-56.12	-51.31	-76.55						
	MeritHigh	-	MeritLow	-3.90	-11.49	-13.68	-22.28	-25.48	-20.04	-10.18	2.21	-13.77						
	GreedHigh	-	GreedLow	3.71	7.33	16.29	15.17	20.46	24.08	27.46	31.00	15.38						
Testing mechanism	RandomHigh	-	MeritHigh	-5.38	-13.48	-25.91	-28.48	-37.20	-64.17*	-79.21*	-83.45*	-110.00**	-120.00**	-128.00**	-146.00**	-157.00**	-164.00**	-174.00*
(high inequality)	RandomHigh	-	GreedHigh	-3.88	-9.15	-16.58	-24.48	-33.70	-45.33	-51.88	-49.45	-67.55	-73.15	-75.24	-90.61	-86.97	-91.15	-90.61
	MeritHigh	-	GreedHigh	1.50	4.33	9.33	4.00	3.50	18.83	27.33	34.00	42.33	47.17	52.33	55.67	70.33	73.17	83.33
Testing mechanism	RandomLow	-	MeritLow	-3.31	-13.87	-20.83*	-25.38	-27.03	-33.50	-33.27	-29.93	-47.11						
(low inequality)	RandomLow	-	GreedLow	5.80	9.29	18.48	16.07	22.41	29.46	31.70	32.86	24.38						
	MeritLow	-	GreedLow	9.11*	23.16**	39.31***	41.45**	49.44**	62.96**	64.97**	62.79*	71.48*						

Note: The table reports differences between the expectations and trustworthiness (measured in tokens) in the treatment reported in the second column and the treatment reported in the third column for low-income subjects below the Merit and Greed thresholds. The treatments either have the same mechanism ('Testing inequality') or the same level of inequality ('Testing mechanism'). ***, ** and * indicates significance at the 1%, 5% and 10% respectively in the Mann-Whitney rank sum test.

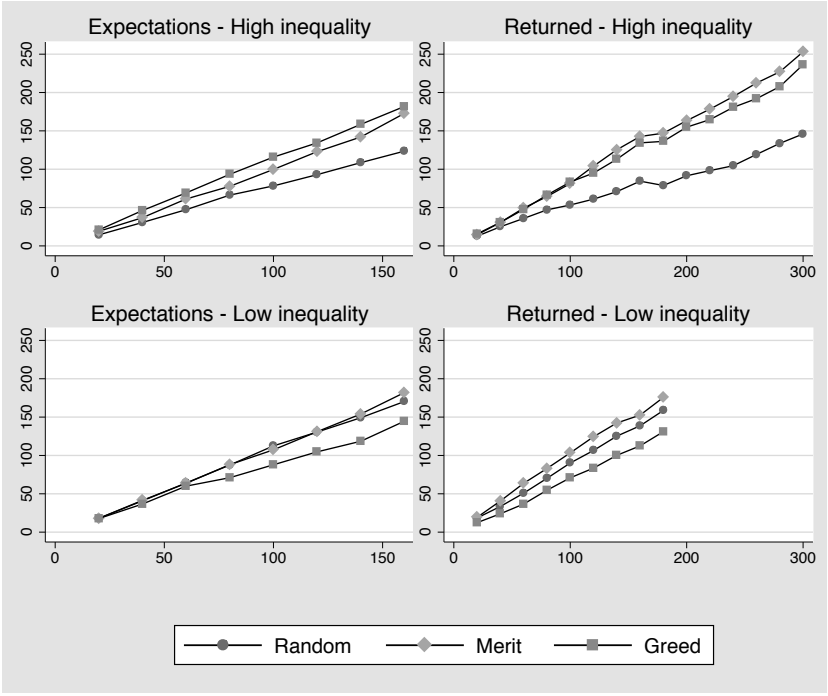
Table 5.B9: Regressions for high-income group

	Trust			Expectations			Trustworthiness		
	(a)	(b)	(c)	(a)	(b)	(c)	(a)	(b)	(c)
Constant	140.00***	238.95***	178.04*	-3.41	78.33	49.91	1.61	44.24	-2.30
	(24.98)	(80.43)	(91.22)	(12.42)	(49.42)	(38.42)	(14.16)	(62.34)	(41.31)
RandomLow	-78.00**	-88.48**	-111.17***	72.70***	62.69***	15.44	67.72***	61.53***	20.86
	(35.33)	(36.40)	(39.72)	(19.70)	(21.34)	(14.95)	(21.59)	(22.29)	(17.13)
MeritHigh	-56.00	-34.13	-36.30	0.26	7.07	5.54	5.72	1.13	-4.40
	(35.33)	(42.94)	(41.95)	(23.69)	(26.88)	(9.93)	(25.94)	(28.21)	(12.76)
MeritLow	-92.00**	-77.37*	-85.00**	21.70	24.80	6.16	29.11	24.22	8.09
	(35.33)	(39.70)	(39.29)	(20.81)	(22.95)	(11.32)	(22.03)	(23.81)	(12.99)
GreedHigh	-62.00*	-48.10	-50.52	-6.19	9.01	6.24	-5.94	4.67	-0.52
	(35.33)	(37.43)	(36.81)	(22.01)	(23.86)	(9.82)	(22.41)	(24.61)	(12.02)
GreedLow	-82.00**	-67.49*	-65.81*	-0.02	15.23	45.75**	-47.28***	-36.96**	-47.01***
	(35.33)	(37.48)	(39.11)	(23.55)	(25.35)	(20.67)	(15.78)	(18.32)	(16.87)
Correct answers in calculation task		-3.07	-3.30		-0.27	-1.45		1.50	1.73
		(4.22)	(4.14)		(2.42)	(1.52)		(2.52)	(1.60)
Tokens kept in decider task		-0.17	-0.07		-0.19*	-0.08		-0.14	-0.00
		(0.15)	(0.15)		(0.10)	(0.06)		(0.13)	(0.08)
Risk aversion			-1.78			-1.72			-1.78
			(6.87)			(3.40)			(2.59)
Trustworthiness			0.07			0.81***			
			(0.19)			(0.07)			
Expectations			0.22						0.69***
			(0.17)						(0.10)
Obs	60	60	60	540	540	540	540	540	540
R-squared	0.14	0.17	0.25	0.40	0.42	0.76	0.41	0.43	0.76
Sent/received amount fixed effects	No	No	No	Yes	Yes	Yes	Yes	Yes	Yes
Clustered standard errors	No	No	No	Yes	Yes	Yes	Yes	Yes	Yes
P-values for Wald test									
Testing inequality									
MeritHigh=MeritLow	0.31	0.24	0.18	0.42	0.48	0.94	0.39	0.37	0.22
GreedHigh=GreedLow	0.57	0.59	0.68	0.82	0.82	0.05	0.02	0.02	0.00
Testing mechanism									
RandomHigh=MeritHigh	0.12	0.43	0.39	0.99	0.79	0.58	0.83	0.97	0.73
RandomHigh=GreedHigh	0.08	0.20	0.18	0.78	0.71	0.53	0.79	0.85	0.97
MeritHigh=GreedHigh	0.87	0.74	0.74	0.81	0.95	0.95	0.67	0.90	0.77
RandomLow=MeritLow	0.69	0.79	0.54	0.03	0.16	0.48	0.09	0.16	0.38
RandomLow=GreedLow	0.91	0.61	0.34	0.01	0.13	0.13	0.00	0.00	0.00
MeritLow=GreedLow	0.78	0.81	0.66	0.41	0.74	0.04	0.00	0.01	0.00

Note: OLS regressions for high-income subjects of trust, expectations and trustworthiness (measured in tokens) on (a) treatments and (b) treatments and performance in Calculation Task and Decider Task. ***, ** and * indicates significance at the 1%, 5% and 10% respectively. Standard errors are given in parentheses.

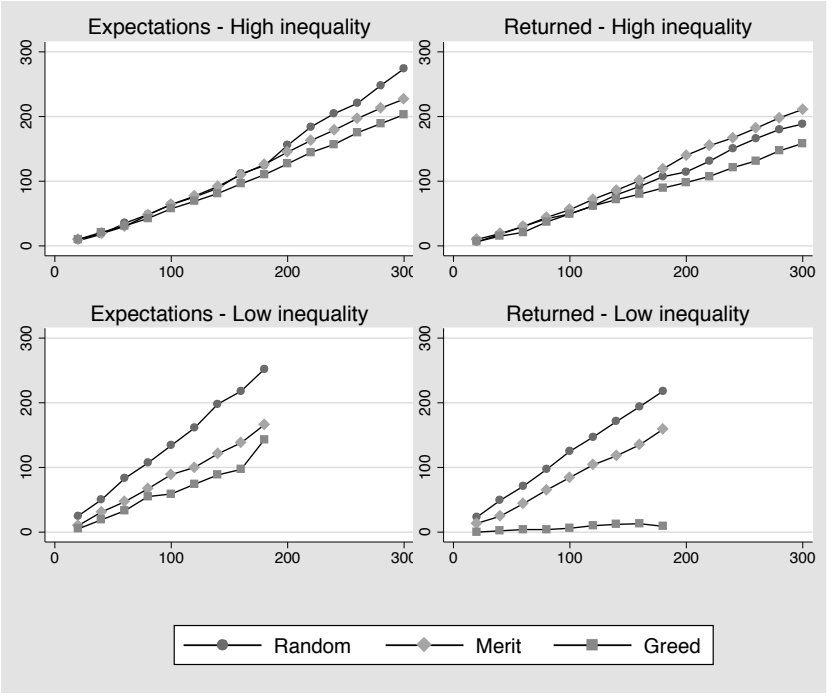
5.C Figures

Figure 5.C1: Expectations and trustworthiness for low-income group by Inequality



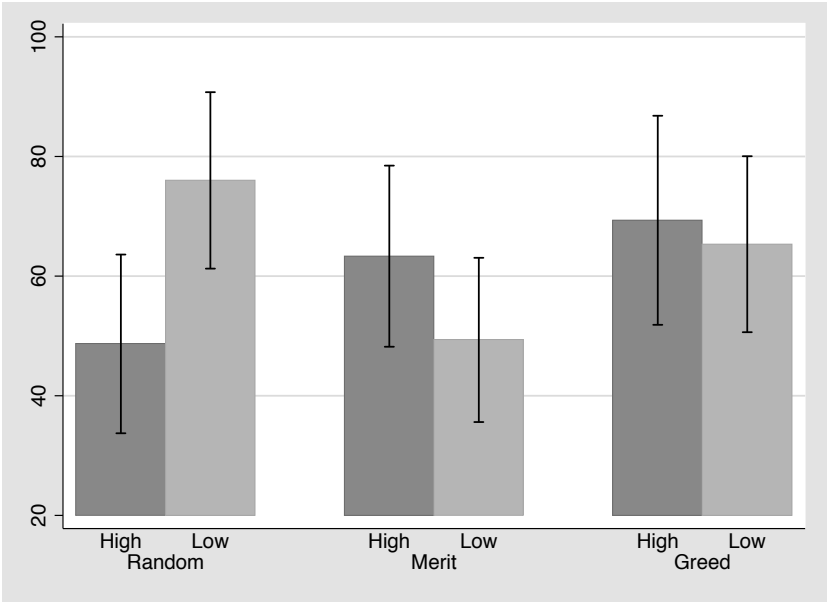
Note: The figure shows the expectations (left column) and trustworthiness (right column) for High and Low inequality treatments respectively for low-income subjects. In the left column, the x-axis indicates the hypothetical amount sent (ranging from 0 to 160 for low-income subjects) and the y-axis indicates the expected amount returned. In the right column, the x-axis indicates the hypothetical amount received (ranging from 0 to 300) and the y-axis indicates the amount returned. Amounts are in tokens.

Figure 5.C2: Expectations and trustworthiness for high-income group by Inequality



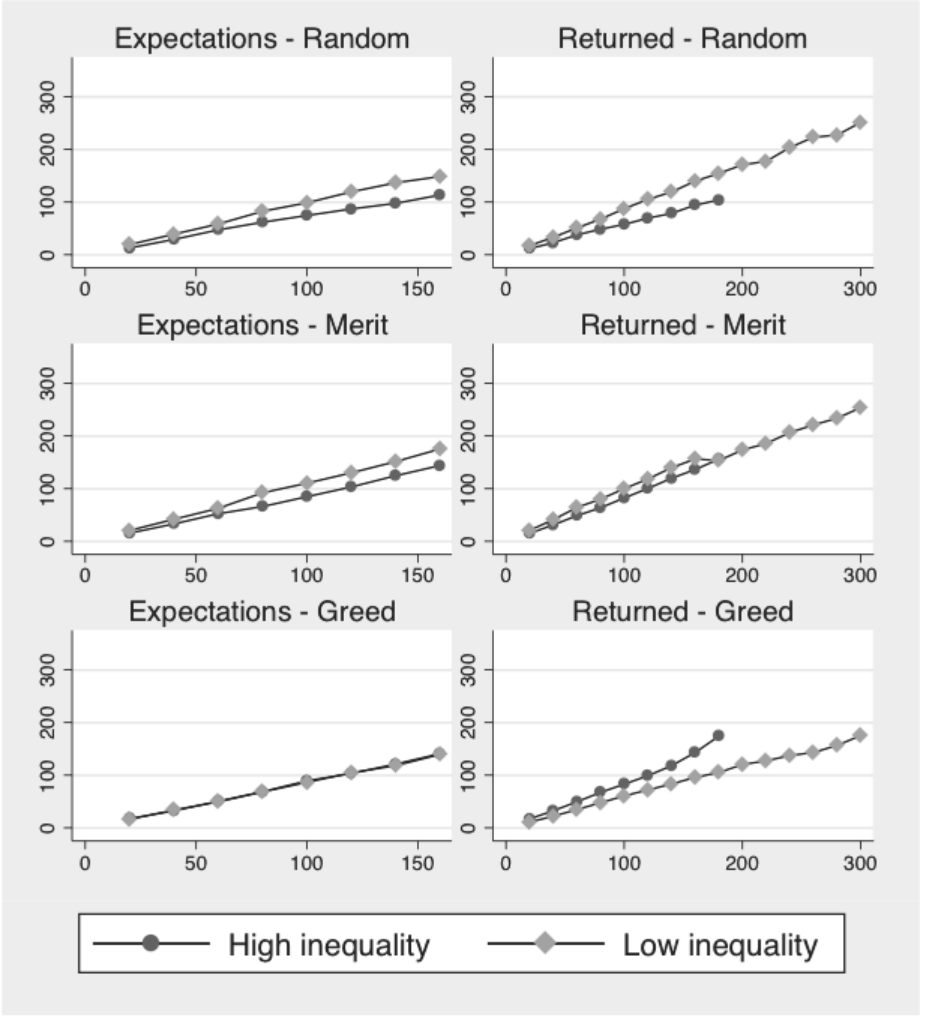
Note: The figure shows the expectations (left column) and trustworthiness (right column) for High and Low inequality treatments respectively for high-income subjects. In the left column, the x-axis indicates the hypothetical amount sent (ranging from 0 to 300 for high-income subjects) and the y-axis indicates the expected amount returned. In the right column, the x-axis indicates the hypothetical amount received (ranging from 0 to 300) and the y-axis indicates the amount returned. Amounts are in tokens.

Figure 5.C3: Trust for low-income group in second period



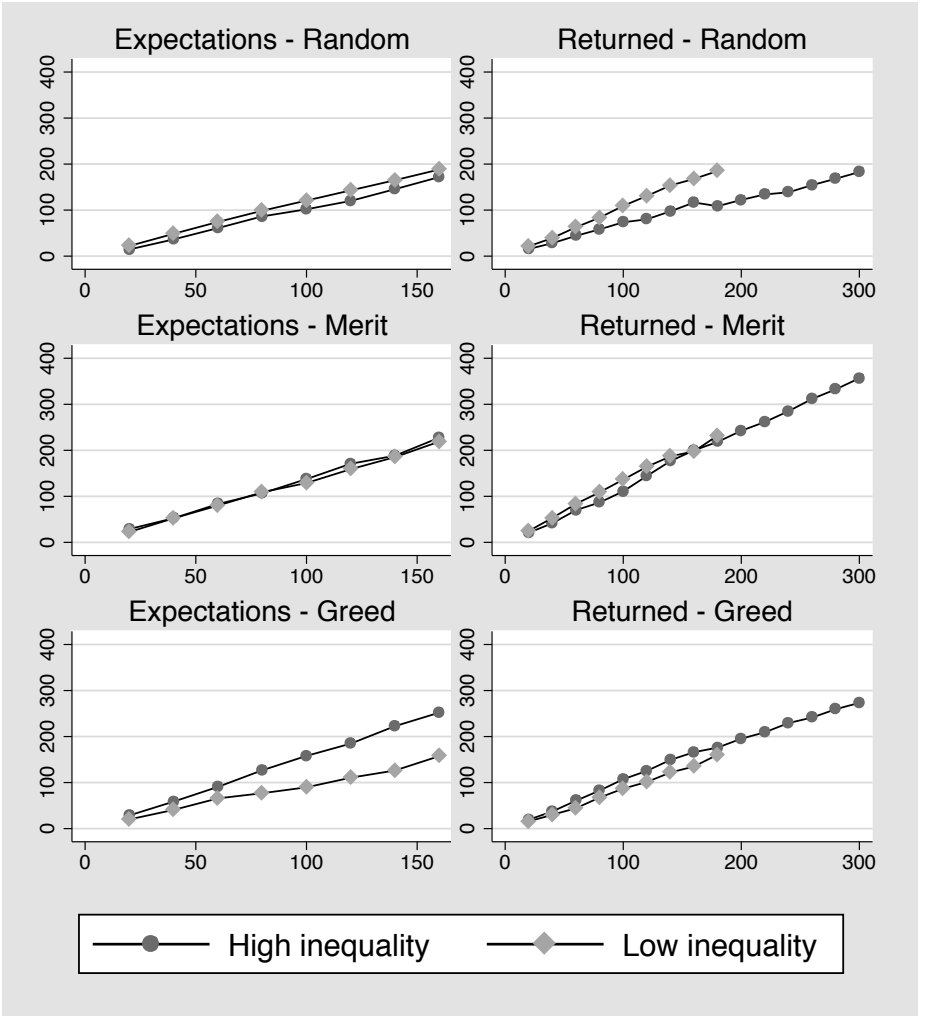
Note: The amount sent in period 2 for low-income subjects in high and low income inequality, according to mechanism. Error bars indicate 10% confidence intervals. Amounts are in tokens.

Figure 5.C4: Expectations and trustworthiness for low-income group by Mechanism in second period



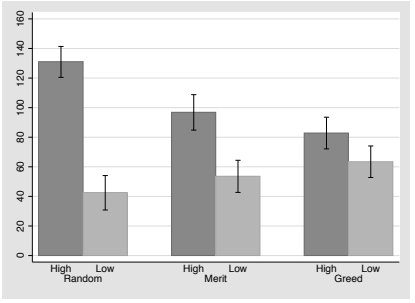
Note: The figure shows the expectations (left column) and trustworthiness (right column) for the Random, Merit and Greed treatment respectively for low-income subjects in period 2. In the left column, the x-axis indicates the hypothetical amount sent (ranging from 0 to 160 for low-income subjects) and the y-axis indicates the expected amount returned. In the right column, the x-axis indicates the hypothetical amount received before tripling (ranging from 0 to 300) and the y-axis indicates the amount returned. Amounts are in tokens.

Figure 5.C5: Expectations and trustworthiness for low-income group by mechanism, corrected for selection effects

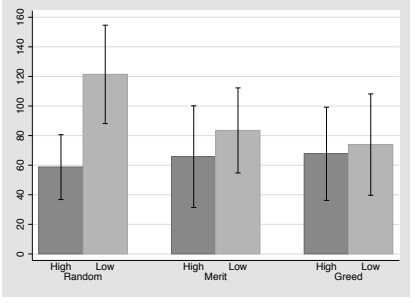


Note: The figure shows the expectations (left column) and trustworthiness (right column) for High and Low inequality treatment respectively for a selected sample of low-income subjects. In the left column, the x-axis indicates the hypothetical amount sent (ranging from 0 to 160 for low-income subjects) and the y-axis indicates the expected amount returned. In the right column, the x-axis indicates the hypothetical amount received before tripling (ranging from 0 to 300) and the y-axis indicates the amount returned. Amounts are in tokens.

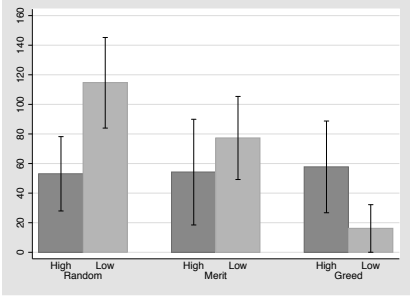
Figure 5.C6: Marginal effects for high-income group, corrected for selection effects parametrically



(a) Trust



(b) Expectations



(c) Trustworthiness

Summary

Chapter 2 (based on joint work with Philip Hans Franses) analyses the effect of mobile phone use on household income in Uganda. Mobile phones have spread rapidly in developing countries over the last decade. It is now widely understood that cell phones have beneficial effects in these areas, but little is known about the impact at the household level.

A two-way causal relationship between economic development and telephone use is not unlikely, making it difficult to obtain unbiased estimates for the impact of mobile phones. After all, someone might become richer as a result of using a cell phone, but cell phone use also relies on having money. We use a novel econometric technique to handle the potential endogeneity in our statistical model. We model the correlation between the endogenous regressor and the error term with copulas, in line with Park and Gupta (2012). To the best of our knowledge, this is the first time that the copula method is applied in the economic development literature.

We find a positive impact of mobile phone use. More specifically, the proportion of mobile phone users in the household as well as the duration of mobile phone ownership matter for household income. We find no evidence for a significant effect of the number of weekly mobile phone calls. Beyond conventional telephone use, we look at mobile banking and mobile search. Very few respondents use these services, but the descriptive statistics indicate that users are on average more economically developed.

In Chapter 3 I move beyond monetary wellbeing and look at life satisfaction (i.e. happiness, subjec-

tive well-being). Developing countries are underrepresented in the happiness literature, which is why I focus on Bolivia. In particular, I study the relationship between happiness and social reference groups. Social reference groups consist of individuals a person interacts with directly and frequently, such as neighbours and colleagues. It is known that individuals compare themselves to people around them and that their relative standing has an impact on their wellbeing. However, current studies rarely analyse social reference groups and rather focus on proxies of reference groups, such as people of the same age and gender.

In this chapter, I analyse self-reported, and thus subjective, comparisons to several reference groups. An obvious complication in studying the relation between two subjective variables - in my case subjective wellbeing and subjective comparisons - is the potential omitted variable bias resulting from a generally positive or negative disposition of respondents. For example, if an individual is naturally optimistic, she will answer more favourably to both subjective questions than a respondent with a more pessimistic outlook. My econometric methodology, similar to Mangyo and Park (2011), deals with this omitted variable bias by estimating a variable that measures disposition. I investigate and confirm the robustness of this method by using anchoring vignettes. The methodology does not deal with potential simultaneous causality of comparisons and life satisfaction. A two-way causal relationship is not unlikely, as happiness impacts factors that are important to relative standing, such as income. I can therefore not draw any conclusions about causality.

My findings support the importance of social reference groups and thereby confirm the relevance of these measures for research dealing with life satisfaction. More specifically, comparisons to family members and (former) classmates are important determinants of subjective wellbeing. I find that other reference groups are associated with future life satisfaction, namely colleagues and neighbours. More generally, I find that the drivers of life satisfaction are in line with present research. The exception is marital status, for which I do not find a significant impact. In addition, I find that future satisfaction differs from present satisfaction, both in terms of distribution and determinants. Overall, I can conclude

that both social reference groups and future life satisfaction provide additional valuable sources of information for subjective wellbeing research.

Chapter 4 focuses on yet another determinant of wellbeing: the happiness of the people around us. Social bonds tend to exist between individuals who are similar with respect to background, characteristics, and behaviour. Do similar people simply seek each other out (*homophily*) or do they also become more similar because of their interaction (*induction*)? This chapter aims to disentangle these two effects, using a unique data set from a Dutch college sorority. The contributions to current literature are twofold. First, it is one of the few network studies related to happiness and, to my knowledge, the first to study happiness in the setting of a fraternal organisation. Second, it is one of the few network studies on fraternal organisations and one of the first quantitative studies on Dutch fraternal organisations.

More specifically, I wish to study the effect of a peer's happiness on an individual's happiness. Statistical models are likely to suffer from endogeneity for several reasons. First, while the peer could affect the individual, it is also likely that the individual in turn influences the peer. Second, the happiness levels among individuals can be correlated for other reasons, such as a similar background or common changes to the environment (*confounding factors*). The setup of the sorority enables me to deal with these issues. Each year a new cohort joins the sorority and new members are expected to form groups within the cohort ('clubs'). Using data from before and after club formation, I am able to disentangle homophily and induction. Data on friendships allow me to control for confounding factors, assuming that friends are more likely to affect each other when they are closer socially.

Next to happiness, I look at other variables of interest, such as body mass index (BMI), alcohol use, and study performance. I do not find evidence for homophily in happiness, behaviour or educational outcomes. However, parental income and existing networks are important in the sorting process: respondents tend to be connected if they come from the same city, both live in a sorority house or follow the same study major. Moreover, I find that happiness, alcohol use, and grades are subject to induction, whereas correlation in BMI and relationship status are due to confounding factors. Checks confirm the

robustness of these effects. Since the induction effects are mainly driven by mutual friendships, they are possibly a consequence of homophily of friendships.

While the preceding two chapters focused on life satisfaction, Chapter 5 (based on joint work with David Smerdon) analyses another 'soft' element of wellbeing: trust. A large body of empirical research has found a negative correlation between trust and income inequality, and it has been suggested that the causal direction runs from inequality to trust. However, little is known about how this relationship depends on the *income distribution mechanism* in a society, which is likely to be relevant. For example, the impact of inequality is probably different in a society where citizens' incomes are based on merit than in a society where greed is rewarded.

Endogeneity is likely to occur in statistical models estimating the impact of inequality on trust. After all, inequality may affect trust, but trust may also determine inequality. Also, unobserved factors, such as cultural traits, could influence both trust and inequality. We therefore run a laboratory experiment in which we randomly assign participants to different scenarios. In our design, individuals are first placed in either a small, high-income class or a larger, low-income class, following a merit-based, greed-based or random allocation. A further treatment variable is the degree of inequality. Subjects then play a trust game against anonymous partners, including the elicitation of expectations with regards to the trustworthiness of their opponent.

Our main findings can be summarised as follows: Income inequality negatively impacts trust in a society in which income classes are determined randomly. When the income distribution mechanism is based on either merit or greed, however, we cannot conclude that changes in income inequality affect trust within the group. Our findings are robust against selection effects. Also, we show that expectations, rather than sending behaviour, are correlated to survey measures of trust. We suggest that our results may be driven by the influence of the distribution mechanism on ingroup/outgroup effects.

Samenvatting (Summary in Dutch)

"Not everything that counts can be counted, and not everything that can be counted counts."
(Niet alles wat telt kan geteld worden, en niet alles wat geteld kan worden telt.)
– Spreuk in Einstein's kantoor op Princeton University

"Hoe gaat het met je?" Het is een vraag die we bijna dagelijks te horen krijgen. Het antwoord lijkt gemakkelijk: "Goed." Echter, het is waarschijnlijk één van de meest complexe vragen die iemand kan stellen. Het vraagt ons immers om alle facetten van ons leven te evalueren: baan, relaties, gezondheid, enzovoorts.

Welzijn staat voor de kwaliteit van leven van een individu of een groep. Economen en beleidsmakers zijn geïnteresseerd in kwantitatieve maatstaven van welzijn. Met andere woorden, ze willen een cijfer geven aan het antwoord op de vraag "Hoe gaat het met je?" Net zoals het niet gemakkelijk is om deze vraag te beantwoorden, is het moeilijk om welzijn te meten.

Economen gebruiken vaak monetaire maatstaven om welzijn te schatten, zoals het bruto nationaal product (BNP) per hoofd van de bevolking. Zulke maatstaven zijn - tenminste in theorie - relatief gemakkelijk te berekenen door de verschillende componenten bij elkaar op te tellen. Zoals het bovenstaande citaat echter duidelijk maakt, kunnen niet alle belangrijke aspecten van ons leven geteld worden. Welzijn wordt beïnvloed door factoren die moeilijk of onmogelijk te tellen zijn, zoals de kwaliteit van onze vriendschappen of hoe we ons voelen als we 's morgens wakker worden.

Het is niet vanzelfsprekend dat mensen die welgesteld zijn ook een hoog welzijn hebben. Daarom zijn economen in toenemende mate geïnteresseerd in non-monetaire maatstaven van welzijn. Een goed voorbeeld is gelukseconomie, een jong vakgebied dat steeds meer erkend wordt als serieus onderzoek. Het Stiglitz-Sen-Fitoussi rapport uit 2009 benadrukte het belang om verder te kijken dan BNP bij het evalueren van kwaliteit van leven. Dit rapport heeft meer belangstelling gecreëerd onder beleidsmakers, waardoor het steeds gebruikelijker wordt dat statistische bureaus data verzamelen over, bijvoorbeeld, geluk.

In dit proefschrift streef ik naar een beter begrip van de factoren die welzijn beïnvloeden. Hoofdstuk 2 meet welzijn op de traditionele manier, door middel van inkomen, terwijl de daaropvolgende hoofdstukken breder kijken door zich te richten op geluk en vertrouwen. Zelfs als er overeenstemming bestaat over hoe welzijn gedefiniëerd en gemeten wordt, is het niet gemakkelijk om het effect van bepaalde factoren te analyseren. Ten eerste, de determinanten van welzijn worden soms op hun beurt beïnvloed door welzijn zelf. Bijvoorbeeld, een hoog inkomen maakt misschien gelukkiger, maar gelukkige mensen hebben ook meer kans op een goede baan met een hoog salaris. Ten tweede, zowel welzijn als zijn invloedsfactoren staan onder invloed van dezelfde onwaarneembare aspecten. Iemand die van nature optimistisch is zal bijvoorbeeld zowel een hoger welzijn als een betere gezondheid hebben.

Deze twee complicaties maken het ingewikkeld om de impact van bepaalde factoren op welzijn te bestuderen. In econometrisch jargon zeggen we dat statistische modellen aan 'endogeniteit' lijden en dat geschatte effecten 'afwijkend' zijn. In alle hoofdstukken krijg ik met endogeniteit te maken en ik ga er op verschillende manieren mee om. In Hoofdstuk 2 en 3 schat ik extra variabelen die - als ik ze meeneem in mijn model - ervoor zorgen dat endogeniteit zo min mogelijk een rol speelt. In Hoofdstuk 4 analyseer ik een groep, een studentenvereniging, met unieke kenmerken die het mogelijk maken om effecten te schatten die niet afwijkend zijn. Hoofdstuk 5 pakt het endogeniteitsprobleem aan op de meest rigoureuze manier, door in een experiment deelnemers willekeurig aan verschillende scenarios

toe te wijzen.

Dit proefschrift is gebaseerd op data die ik zelf heb verzameld in Uganda, Bolivia en Nederland. Als econometrist gebruik ik kwantitatieve data om relevante onderzoeksvragen te beantwoorden. Omdat ik mijn respondenten zelf heb gesproken, heb ik echter ook veel geleerd over de waarde van kwalitatieve data. Naast het beantwoorden van mijn vragen, lichtten respondenten hun antwoord vaak toe of ze vertelden me meer over hun leven in het algemeen. Zulke informatie is belangrijk, zeker als het om concepten gaat die moeilijk te kwantificeren zijn, zoals welzijn. Ik ben er daarom van overtuigd dat welzijnsonderzoek baat heeft bij een combinatie van kwantitatieve en kwalitatieve methoden. Hoewel ik er niet voor opgeleid ben om kwalitatief onderzoek te doen, wilde ik mijn respondenten toch een stem geven naast de gemiddeldes, coëfficiënten en standaardfouten. Elk hoofdstuk begint om die reden met een citaat van iemand die aan mijn onderzoek heeft deelgenomen.

Overzicht

Hoofdstuk 2 (in samenwerking met Philip Hans Franses) onderzoekt het effect van mobiele telefoons op huishoudelijk inkomen in Uganda. Mobiele telefoongebruik heeft zich snel verspreid in ontwikkelingslanden en het is bekend dat dit een positief effect heeft gehad op economische ontwikkeling. Echter, we weten nog maar weinig over het effect op huishoudniveau.

Het is waarschijnlijk dat het causaal verband tussen ontwikkeling en telefoons in twee richtingen loopt. Iemand wordt misschien rijker van het telefoongebruik, maar hij of zij heeft in de eerste plaats geld nodig om een telefoon aan te schaffen en te bellen. Hierdoor is het moeilijk om een goede schatting te krijgen van het effect van de mobiele telefoon. Om die reden gebruiken we een nieuwe econometrische techniek om met de endogeniteit in ons model om te gaan. Aan de hand van Park and Gupta (2012) modelleren we de correlatie tussen de endogene regressor en de standaardfout door middel van ‘copulas.’ Voor zover ons bekend is, is dit de eerste keer dat de copula methode in de

ontwikkelingseconomie wordt toegepast.

We vinden een positief effect van mobiel telefoongebruik. Zowel de proportie van telefoongebruikers in het huishouden als de duur van telefoonbezit spelen een belangrijke rol. We vinden geen aanwijzingen voor een effect van de frequentie van telefoongesprekken. Daarnaast bestuderen we het gebruik van mobiel bankieren en mobiele zoekservices. Zeer weinig respondenten maken hier gebruik van, maar de beschrijvende statistieken laten zien dat gebruikers gemiddeld meer economisch ontwikkeld zijn.

In Hoofdstuk 3 kijk ik breder dan monetair welzijn en richt ik me op geluk. Ontwikkelingslanden zijn ondervertegenwoordigd in de geluksliteratuur. Om die reden bestudeer ik Bolivia, waar ik het verband tussen geluk en sociale referentiegroepen onderzoek. Sociale referentiegroepen bestaan uit mensen waar een persoon regelmatig mee in contact komt, zoals buren of collega's. Het is bekend dat individuen zich vergelijken met de mensen om hen heen and dat hun relatieve positie bepalend is voor hun welzijn. Huidige studies kijken echter zelden naar sociale referentiegroepen en gebruiken eerder benaderingen, zoals mensen van hetzelfde geslacht met dezelfde leeftijd.

In dit hoofdstuk analyseer ik zelfgerapporteerde, en dus subjectieve, vergelijkingen met verscheidene referentiegroepen. Omdat ik het verband tussen twee subjectieve variabelen bestudeer - geluk en vergelijkingen - moet ik ermee rekening houden dat bepaalde onwaarneembare factoren de relatie kunnen beïnvloeden. Als iemand namelijk van nature optimistisch is, zal hij of zij beide subjectieve vragen positiever beantwoorden dan een pessimist. Mijn econometrische methode, gebaseerd op Mangyo and Park (2011), schat de mate van optimisme per persoon. Door middel van 'anchoring vignettes' - vragen over het geluk van hypothetische personen - laat ik zien dat mijn methode inderdaad een schatting geeft van optimisme. Mijn analyse is ook gevoelig voor het mogelijke effect van geluk op vergelijkingen. Een dergelijk effect is niet onwaarschijnlijk, omdat geluk factoren beïnvloedt die belangrijk zijn voor de vergelijking met de omgeving, zoals inkomen. Om die reden kan ik geen conclusies trekken over de causaliteit van mijn bevindingen.

Mijn resultaten bevestigen het belang van sociale referentiegroepen voor geluk. Vergelijkingen met familieleden en (voormalige) klasgenoten correleren met geluk, terwijl collega's en buren belangrijk zijn voor toekomstig geluk. Verder vind ik dat overige invloedsfactoren overeenkomen met de geluksliteratuur. De uitzondering is burgerlijke staat, waarvoor ik geen significant effect vind. Bovendien verschillen zowel de verdeling als de determinanten van toekomstig geluk van die van huidig geluk.

Hoofdstuk 4 beschouwt een andere determinant van welzijn: het geluk van de mensen om ons heen. Sociale contacten bestaan voornamelijk tussen individuen die op elkaar lijken qua achtergrond, eigenschappen of gedrag. Zoeken gelijkaardige mensen elkaar simpelweg op (*homofilie*) of gaan ze ook meer op elkaar lijken als gevolg van hun contact (*inductie* of besmetting)? Dit hoofdstuk ontrafelt deze twee effecten door een studentenvereniging te bestuderen. Het is één van de eerste netwerkstudies over geluk en, voorzover mij bekend, de eerste die geluk bestudeerd in een studentenvereniging. Bovendien is het één van de weinige netwerkstudies over studentenverenigingen en slechts de tweede kwantitatieve analyse van een Nederlandse studentenvereniging.

Ik analyseer het effect van het geluk van een ander op het eigen geluk. Statistische modellen krijgen mogelijk te maken met endogeniteit. Ten eerste is het aannemelijk dat een persoon niet alleen beïnvloed wordt door een ander, maar dat zij zelf ook een effect heeft op de mensen om hem heen. Ten tweede is het geluk van personen mogelijk gecorreleerd om andere redenen, zoals een soortgelijke achtergrond of de ervaring van dezelfde gebeurtenissen (*verstorende factoren*). De structuur van de vereniging maakt het mogelijk om met deze moeilijkheden om te gaan. Elk jaar voegt een nieuwe lichting zich bij de vereniging en deze nieuwe leden vormen jaarclubs, groepen binnen de lichting. Ik kan inductie en homofilie onderscheiden door data te gebruiken van vóór en na de clubvorming. Het effect van verstorende factoren wordt getoetst met data over vriendschappen. Hierbij neem ik aan dat personen elkaar meer beïvloeden als ze een hechtere band hebben.

Naast geluk bestudeer ik ook andere variabelen, zoals *body mass index* (BMI), alcoholgebruik en studieprestaties. Ik vind geen aanwijzingen voor homofilie van geluk, gedrag of studie uitkomsten.

Het inkomen van ouders en bestaande netwerken zijn wel belangrijk: respondenten zijn eerder verbonden als hun ouders rijk zijn, ze uit dezelfde woonplaats komen, beide in een verenigingshuis wonen of dezelfde studie doen. Verder vind ik dat geluk, alcoholgebruik en studiecijfers onderhevig zijn aan inductie, terwijl correlaties in BMI en relaties worden veroorzaakt door verstorende factoren. De inductie-effecten kunnen voornamelijk verklaard worden door wederzijdse vriendschappen. Ik kan daarom niet uitsluiten dat deze effecten een gevolg zijn van homofilie in vriendschapsvorming.

Terwijl de voorgaande twee hoofdstukken over geluk gingen, analyseert Hoofdstuk 5 (in samenwerking met David Smerdon) een ander ongrijpbaar element van welzijn: vertrouwen. Een grote hoeveelheid empirische studies heeft een negatieve correlatie gevonden tussen inkomensongelijkheid en vertrouwen. Zulke artikelen suggereren dat ongelijkheid vertrouwen beïnvloedt in plaats van andersom. Toch is weinig bekend over de causaliteit van de relatie en hoe het verband bepaald wordt door verschillende *inkomensverdelingsmechanismen*. De invloed van ongelijkheid is namelijk waarschijnlijk anders in een samenleving waar inkomen wordt verdeeld op basis van verdienste dan, bijvoorbeeld, in een samenleving waar hebzucht wordt beloond.

Endogeniteit speelt waarschijnlijk een rol in statistische modellen die het effect van ongelijkheid op vertrouwen analyseren. Ongelijkheid kan immers vertrouwen beïnvloeden, maar vertrouwen kan ook een impact hebben op ongelijkheid. Bovendien kunnen onwaarneembare factoren, zoals culturele aspecten, van invloed zijn op zowel ongelijkheid als vertrouwen. Om die reden doen we een computerexperiment waarin we individuen willekeurig toewijzen aan een bepaald scenario. Op deze manier is er geen zelfselectie mogelijk en weten we zeker dat vertrouwen geen invloed kan hebben op inkomensongelijkheid. Deelnemers worden eerst opgedeeld in een rijke, kleine, inkomensgroep of een arme, grote, inkomensgroep. Deze verdeling wordt gemaakt op basis van verdienste, aan de hand van hebzucht of compleet willekeurig. Vervolgens krijgen de deelnemers te maken met óf hoge óf lage ongelijkheid. Ze spelen een ‘vertrouwensspel’ met een anonieme tegenspeler, waarin we deelnemers ook vragen om hun verwachtingen aan te geven.

We vinden dat hogere inkomensongelijkheid lager vertrouwen veroorzaakt, maar alleen als inkomen willekeurig wordt toegewezen. We vinden geen verschil tussen hoge en lage ongelijkheid als we inkomen verdelen op basis van verdienste of hebzucht. Onze bevindingen worden niet beïnvloed door selectie-effecten. We geven *ingroup/outgroup* effecten als een mogelijke verklaring van onze resultaten.

Bibliography

Abraham, R. (2007). "Mobile phones and economic development: Evidence from the fishing industry in India." *Information Technologies and International Development*, 4(1), 5–17.

Aker, J. (2008). "Does digital divide or provide? The impact of cell phones on grain markets in Niger." *Working Paper 154*, Center for Global Development.

Alesina, A. and La Ferrara, E. (2002). "Who trusts others?" *Journal of Public Economics*, 85(2), 207–234.

Anderson, L., Mellor, J., and Milyo, J. (2006). "Induced heterogeneity in trust experiments." *Experimental Economics*, 9(3), 223–235.

Angrist, J. D. and Lang, K. (2004). "Does school integration generate peer effects? Evidence from Boston's Metco program." *American Economic Review*, 94(5), 1613–1634.

Arrow, K. (1974). *The Limits of Organization*. Fels Center of Government Series. W W Norton & Company Incorporated.

Averett, S., Terrizzi, S., and Wang, Y. (2013). "The effect of sorority membership on eating disorders and body mass index." *Discussion Paper 7512*, Institute for the Study of Labor (IZA).

Berg, J., Dickhaut, J., and McCabe, K. (1995). "Trust, reciprocity, and social history." *Games and Economic Behavior*, 10(1), 122–142.

Blanchflower, D. G. and Oswald, A. J. (2004). "Well-being over time in Britain and the USA." *Journal of Public Economics*, 88(7-8), 1359–1386.

Blanchflower, D. G. and Oswald, A. J. (2008). "Is well-being u-shaped over the life cycle?" *Social Science & Medicine*, 66(8), 1733–1749.

Brülhart, M. and Usunier, J.-C. (2012). "Does the trust game measure trust?" *Economics Letters*, 115(1), 20–23.

Cantril, H. (1965). *The Pattern of Human Concerns*. Rutgers University Press, New Brunswick, N. J.

Cappelen, A. W., Hole, A. D., Sorensen, E. O., and Tungodden, B. (2007). "The pluralism of fairness ideals: An experimental approach." *American Economic Review*, 97(3), 818–827.

Central Intelligence Agency (2009). "The world factbook 2009." , Washington, DC: CIA.

Chen, S., Schreiner, M., and Woller, G. (2008). "Progress out of poverty index^{tm}: A simple poverty scorecard for Kenya." , Grameen Foundation.

Chen, Y. and Li, S. X. (2009). "Group identity and social preferences." *American Economic Review*, 99(1), 431–57.

Cherry, T. L., Frykblom, P., and Shogren, J. F. (2002). "Hardnose the dictator." *American Economic Review*, 92(4), 1218–1221.

Chetty, R., Friedman, J. N., Hilger, N., Saez, E., Schanzenbach, D. W., and Yagan, D. (2011). "How does your kindergarten classroom affect your earnings? Evidence from project STAR." *The Quarterly Journal of Economics*, 126(4), 1593–1660.

Christakis, N. A. and Fowler, J. H. (2013). "Social contagion theory: Examining dynamic social networks and human behavior." *Statistics in Medicine*, 32(4), 556–577.

Ciriolo, E. (2007). "Inequity aversion and trustees' reciprocity in the trust game." *European Journal of Political Economy*, 23(4), 1007–1024.

Clark, A., Frijters, P., and Shields, M. (2008). "Relative income, happiness, and utility: An explanation for the Easterlin paradox and other puzzles." *Journal of Economic Literature*, 46(1), 95–144.

Clark, A. E. and D'Ambrosio, C. (2014). "Attitudes to income inequality: Experimental and survey evidence." , PSE.

Clark, A. E. and Senik, C. (2010). "Who compares to whom? The anatomy of income comparisons in Europe." *The Economic Journal*, 120(544), 573–594.

Cohen-Cole, E. and Fletcher, J. M. (2008). "Is obesity contagious? Social networks vs. environmental factors in the obesity epidemic." *Journal of Health Economics*, 27(5), 1382–1387.

Coyle, D. (2007). "Overview." *Policy Paper 6, 2-5*, Vodafone Group.

D'Agostino, R. B., Belanger, A. J., and D'Agostino, R. B. (1990). "A suggestion for using powerful and informative tests of normality." *American Statistician*, 44, 316–321.

De Neve, J.-E., Diener, E., Tay, L., and Xuereb, C. (2013). "The objective benefits of subjective well-being." *World Happiness Report 2013*, J. Helliwell, R. Layard, and J. Sachs, eds., New York: UN Sustainable Development Solutions Network.

DeBard, R., Lake, T., and Binder, R. S. (2006). "Greeks and grades: The first-year experience." *NASPA Journal*, 43(1), 56–68.

DeSimone, J. (2007). "Fraternity membership and binge drinking." *Journal of Health Economics*, 26(5), 950–967.

DeSimone, J. (2009). "Fraternity membership and drinking behavior." *Economic Inquiry*, 47(2), 337 – 350.

Dolan, P., Peasgood, T., and White, M. (2008). "Do we really know what makes us happy? A review of the economic literature on the factors associated with subjective well-being." *Journal of Economic Psychology*, 29(1), 94–122.

Donner, J. (2008). "Research approaches to mobile use in the developing world: A review of the literature." *The Information Society*, 24, 140–159.

Dronkers, J. and Hillege, S. M. M. (1998). "Board membership of traditional male fraternities and access to dutch elites: A disappearing avenue to elite positions?" *European Sociological Review*, 14(2), 191–203.

Easterlin, R. (2005a). "Diminishing marginal utility of income? Caveat emptor." *Social Indicators Research*, 70(3), 243–255.

Easterlin, R. A. (1974). "Does economic growth improve the human lot? Some empirical evidence." *Nations and Households in Economic Growth: Essays in Honor of Moses Abramowitz*, P. A. David and M. W. Reder, eds., Academic Press.

Easterlin, R. A. (1995). "Will raising the incomes of all increase the happiness of all?" *Journal of Economic Behavior and Organization*, 27(1), 35–47.

Easterlin, R. A. (2005b). "Feeding the illusion of growth and happiness: A reply to Hagerty and Veenhoven." *Social Indicators Research*, 74(3), 429–443.

Easterlin, R. A. and Sawangfa, O. (2007). "Happiness and domain satisfaction: Theory and evidence." *IZA Discussion Papers 2584*, Institute for the Study of Labor (IZA) (January).

Easterly, W. and Levine, R. (2002). "Tropics, germs and crops: How endowments influence economic development." *Journal of Monetary Economics*, 50, 3–39.

Eccles, M. (1966). *Beckoning Frontiers: Public and Personal Recollections*. Knopf.

Eggleston, K. N., Jensen, R., and Zeckhauser, R. (2002). "Information and communication technologies, markets and economic development." *Discussion Paper 0203*, Department of Economics, Tufts University.

Eibner, C. and Evans, W. (2005). "Relative deprivation, poor health habits, and mortality." *Journal of Human Resources*, 40(3), 591–620.

Emwanu, T., Okwi, P. O., Hoogeveen, J., Kristjanson, P., and Henninger, N. (2007). "Nature, distribution and evolution of poverty and inequality in Uganda." , International Livestock Research Institute.

Fafchamps, M. and Gubert, F. (2007). "The formation of risk sharing networks." *Journal of Development Economics*, 83(2), 326–350.

Fafchamps, M. and Shilpi, F. (2008). "Subjective welfare, isolation, and relative consumption." *Journal of Development Economics*, 86(1), 43 – 60.

Fischbacher, U. (2007). "Z-tree: Zurich toolbox for ready-made economic experiments." *Experimental Economics*, 10(2), 171–178.

Fowler, J. and Christakis, N. (2008). "The dynamic spread of happiness in a large social network." *BMJ: British medical journal*, 337.

Fukuyama, F. (1995). *Trust: The Social Virtues and the Creation of Prosperity*. Free Press, New York.

Gambetta, D. (2000). "Can we trust trust?" *Trust: Making and Breaking Cooperative Relations*, D. Gambetta, ed., University of Oxford, 213–227.

Gilbert, N. and Meyer, C. (2004). "Similarity in young women's eating attitudes: Self-selected versus artificially constructed groups." *International Journal of Eating Disorders*, 36(2), 213–219.

Graham, C. and Pettinato, S. (2002). "Frustrated achievers: Winners, losers and subjective well-being in new market economies." *Journal of Development Studies*, 38(4), 100–140.

Green, M. C. and Brock, T. C. (2005). "Organizational membership versus informal interaction: Contributions to skills and perceptions that build social capital." *Political Psychology*, 26(1), 1–25.

Greiner, B., Ockenfels, A., and Werner, P. (2012). "The dynamic interplay of inequality and trust - an experimental study." *Journal of Economic Behavior & Organization*, 81(2), 355–365.

Grubb, F. (2006). "Does going Greek impair undergraduate academic performance? A case study." *American Journal of Economics and Sociology*, 65(5), 1085–1110.

Gruijter, D. d. et al. (2005). "Over verenigingslidmaatschap en studiesucces." *Internal report 171*, ICLON, Leiden.

GSMA (2013). "Sub-Saharan Africa mobile economy 2013." , GSM Association.

Guryan, J., Kroft, K., and Notowidigdo, M. J. (2009). "Peer effects in the workplace: Evidence from random groupings in professional golf tournaments." *American Economic Journal: Applied Economics*, 1(4), 34–68.

Gustavsson, M. and Jordahl, H. (2008). "Inequality and trust in Sweden: Some inequalities are more harmful than others." *Journal of Public Economics*, 92(1-2), 348–365.

Hardy, A. P. (1980). "The role of the telephone in economic development." *Telecommunications Policy*, 4(4), 278 – 286.

Hargreaves Heap, S., Tan, J., and Zizzo, D. (2013). "Trust, inequality and the market." *Theory and Decision*, 74(3), 311–333.

Hayo, B. and Seifert, W. (2003). "Subjective economic well-being in Eastern Europe." *Journal of Economic Psychology*, 24(3), 329–348.

Heckman, J. (1978). "Dummy endogenous variables in a simultaneous equation system." *Econometrica*, 46(4), 931–959.

Heemskerk, E. and Fennema, M. (2009). "Network dynamics of the Dutch business elite." *International Sociology*, 24(6), 807–832.

Helliwell, J., Barrington-Leigh, C., Harris, A., and Huang, H. (2010). "International evidence on the social context of well-being." *International Differences in Well-Being*, E. Diener, J. Helliwell, and D. Kahneman, eds., Oxford University Press, 213–229.

Helliwell, J., Layard, R., and Sachs, J. D. (2012). *World Happiness Report*. Earth Institute, Columbia University, New York.

Helliwell, J. F. (2003). "How's life? Combining individual and national variables to explain subjective well-being." *Economic Modelling*, 20(2), 331–360.

Hoffman, E., McCabe, K., Shachat, K., and Smith, V. (1994). "Preferences, property rights, and anonymity in bargaining games." *Games and Economic Behavior*, 7(3), 346–380.

Holt, C. and Laury, S. (2002). "Risk aversion and incentive effects." *The American Economic Review*, 92(5), 1644–1655.

Hopkins, D. J. and King, G. (2010). "Improving anchoring vignettes: Designing surveys to correct interpersonal incomparability." *Public Opinion Quarterly*, 74(2), 201–222.

Hufman, P., Straub, P. W., Freund, A., and Sanders, T. N. (2007). *Studentenverenigingen van 1800 tot nu*. Utrecht: Landelijke Kamer van Verenigingen.

Ivatury, G. and Mas, I. (2008). "The early experience with branchless banking." , Washington DC: Consultative group to assist the poor (CGAP) and the United Nations Foundation.

Ivatury, G. and Pickens, M. (2002). "Mobile phone banking and low-income countries: Evidence from South Africa." *International Journal of Retail & Distribution Management*, 30(6), 331–339.

Iwashyna, T., Hutchins, M., Gebremariam, C., and Lee, J. (2011). "An agent-based approach to evaluating the performance characteristics of regression for distinguishing longitudinal network effects: Pragmatic lessons.

Jarque, C. M. and Bera, A. K. (1987). "A test for normality of observations and regression residuals." *International Statistical Review*, 2, 163–172.

Jensen, R. (2007). "The digital provide: Information (technology), market performance, and welfare in the South Indian fisheries sector." *The Quarterly Journal of Economics*, 122(3), 879–924.

Johnson, N. D. and Mislin, A. (2012). "How much should we trust the World Values Survey trust question?" *Economics Letters*, 116(2), 210–212.

Jones, M., Buchanan, G., Harpe, R., and Xech, P. (2007). "Questions not answers: A novel mobile search technique." *CHI '07: Proceedings of the SIGCHI conference on Human factors in computing systems*, New York, NY, USA, ACM, 155–158.

Jordahl, H. (2008). "Economic inequality." *Handbook of Social Capital*, G. Svendsen and G. Svendsen, eds., Edward Elgar, Cheltenham, 323–336.

Kamvar, M. and Baluja, S. (2006). "A large scale study of wireless search behavior: Google mobile search." *CHI '06: Proceedings of the SIGCHI conference on Human Factors in computing systems*, New York, NY, USA, ACM, 701–709.

Kapteyn, A., Smith, J. P., and Soest, A. v. (2009). "Life satisfaction." *IZA Discussion Paper Series*, 4015.

Kathuria, R., Uppal, M., and Mamta (2009). "An econometric analysis of the impact of mobile." *Policy Paper 9, 5-20*, Vodafone Group.

Kerkvliet, J. (1994). "Cheating by economics students: A comparison of survey results." *Journal of Economic Education*, 25(2), 121–133.

King, G., Murray, C., Salomon, J., and Tandon, A. (2004). "Enhancing the validity and cross-cultural comparability of measurement in survey research." *American Political Science Review*, 98, 191–207.

Kling, J. R., Liebman, J. B., and Katz, L. F. (2007). "Experimental analysis of neighborhood effects." *Econometrica*, 75(1), 83–119.

Knight, J. and Gunatilaka, R. (2011). "Does economic growth raise happiness in China?" *Oxford Development Studies*, 39(01), 1–24.

Knight, J., Song, L., and Gunatilaka, R. (2009). "Subjective well-being and its determinants in rural China." *China Economic Review*, 20(4), 635 – 649.

Konow, J. (1996). "A positive theory of economic fairness." *Journal of Economic Behavior & Organization*, 31(1), 13–35.

Konow, J. (2000). "Fair shares: Accountability and cognitive dissonance in allocation decisions." *American Economic Review*, 90(4), 1072–1091.

Kramer, R. (1996). "Divergent realities and convergent disappointments in the hierarchic relation: Trust and the intuitive auditor at work." *Trust in organizations: Frontiers of theory and research*, R. Kramer and T. Tyler, eds., SAGE Publications, Inc., Thousand Oaks.

Krueger, J. I., Massey, A. L., and DiDonato, T. E. (2008). "A matter of trust: From social preferences to the strategic adherence to social norms." *Negotiation and Conflict Management Research*, 1(1), 31–52.

Ku, H. and Salmon, T. C. (2013). "Procedural fairness and the tolerance for income inequality." *European Economic Review*, 64(0), 111 – 128.

Kugler, T., Connolly, T., and Kausel, E. E. (2009). "The effect of consequential thinking on trust game behavior." *Journal of Behavioral Decision Making*, 22(2), 101–119.

Laforet, S. and Li, X. (2005). "Consumers' attitudes towards online and mobile banking in China." *International Journal of Bank Marketing*, 23(5), 362–380.

Lee, K. H. and Bellemare, M. F. (2012). "Look who's talking: The impacts of the intrahousehold allocation of mobile phones on agricultural prices." *The Journal of Development Studies*, 49(5), 624–640.

Lei, V. and Vesely, F. (2010). "In-group versus out-group trust: The impact of income inequality." *Southern Economic Journal*, 76(4), 1049–1063.

Leigh, A. (2006a). "Does equality lead to fraternity?." *Economics Letters*, 93(1), 121–125.

Leigh, A. (2006b). "Trust, inequality and ethnic heterogeneity." *The Economic Record*, 82(258), 268–280.

Maalsté, N. (2000). "Ad fundum! Een blik in de gevarieerde drinkcultuur van het Nederlandse studentenleven." , Centrum voor Verslavingsonderzoek, Universiteit Utrecht.

MacKerron, G. (2011). "Happiness economics from 35 000 feet." *Journal of Economic Surveys*, 705–735.

Mangyo, E. and Park, A. (2011). "Relative deprivation and health: Which reference groups matter?" *Journal of Human Resources*, 46(3), 459–481.

Marmaros, D. and Sacerdote, B. (2002). "Peer and social networks in job search." *European Economic Review*, 46(4-5), 870 – 879.

Martin, G. L., Hevel, M. S., and Pascarella, E. T. (2012). "Do fraternities and sororities enhance socially responsible leadership?" *Journal of Student Affairs Research and Practice*, 49(3), 267–284.

Mbiti, I. and Weil, D. N. (2011). "Mobile banking: The impact of M-Pesa in Kenya." *NBER Working Paper Series*.

Mittal, S., Gandhi, S., and Tripathi, G. (2009). "The impact of mobiles on agricultural productivity." *Policy Paper 9, 21-33*, Vodafone Group.

Muto, M. and Yamano, T. (2009). "The impact of mobile phone coverage expansion on market participation: Panel data evidence from Uganda." *World Development*, 37(12), 1887–1896.

Nelsen, R. (2006). *An Introduction to Copulas*. Springer, New York, 2nd edition edition.

Nieuwenhuis, M. and Postmes, T. (2011). "Grensoverschrijdend gedrag onder Groningse studenten." , Instituut voor integrate en sociale weerbaarheid.

Ostry, M. J. D., Berg, M. A., and Tsangarides, M. C. G. (2014). *Redistribution, Inequality, and Growth*. International Monetary Fund.

Pagan, A. (1984). "Econometric issues in the analysis of regressions with generated regressors." *International Economic Review*, 25(1), pp. 221–247.

Park, S. and Gupta, S. (2012). "Handling endogenous regressors by joint estimation using copulas." *Marketing Science*, 31(4), 567–586.

Pascarella, E. T., Flowers, L., and Whitt, E. J. (2001). "Cognitive effects of Greek affiliation in college: Additional evidence." *NASPA Journal*, 38(3), 280–301.

Phua, J. (2011). "The influence of peer norms and popularity on smoking and drinking behavior among college fraternity members: A social network analysis." *Social Influence*, 6(3), 153–168.

Pickett, K. and Wilkinson, R. (2009). *The Spirit Level: Why Greater Equality Makes Societies Stronger*. Bloomsbury Publishing.

Pike, G. R. (2000). "The influence of fraternity or sorority membership on students' college experiences and cognitive development." *Research in Higher Education*, 41(1), 117–139.

Platow, M. J., Foddy, M., Yamagishi, T., Lim, L., and Chow, A. (2012). "Two experimental tests of trust in in-group strangers: The moderating role of common knowledge of group membership." *European Journal of Social Psychology*, 42(1), 30–35.

Porteous, D. (2007). "Just how transformational is m-banking?" , FinMark Trust.

Ravallion, M. (2012). "Poor, or just feeling poor? On using subjective data in measuring poverty." *5968*, World Bank.

Reifman, A., Watson, W. K., and McCourt, A. (2006). "Social networks and college drinking: Probing processes of social influence and selection." *Personality and Social Psychology Bulletin*, 32(6), 820–832.

Royston, P. (1991). "sg3.5: Comment on sg3.4 and an improved D'Agostino test." *Stata Technical Bulletin*, 3, 23–24.

Sacerdote, B. (2001). "Peer effects with random assignment: Results for Dartmouth roommates." *Quarterly Journal of Economics*, 116(2), 681–704.

Sacks, D. W., Stevenson, B., and Wolfers, J. (2010). "Subjective well-being, income, economic development and growth." *Working Paper 16441*, National Bureau of Economic Research (October).

Sapienza, P., Toldra-Simats, A., and Zingales, L. (2013). "Understanding trust." *The Economic Journal*, 1313–1332.

Schreiner, M. (2009). "A simple poverty scorecard for Bolivia." , Microfinance Risk Management, L.L.C.

Scott-Sheldon, L., Carey, K., and Carey, M. (2008). "Health behavior and college students: Does Greek affiliation matter?" *Journal of Behavioral Medicine*, 31(1), 61–70.

Shalizi, C. R. and Thomas, A. C. (2011). "Homophily and contagion are generically confounded in observational social network studies." *Sociological Methods & Research*, 40(2), 211–239.

Shiller, R. J., Boycko, M., and Korobov, V. (1991). "Popular attitudes toward free markets: The Soviet Union and the United States compared." *American Economic Review*, 81(3), pp. 385–400.

Sklar, A. (1973). "Random variables, joint distribution functions, and copulas." *Kybernetika*, 9(6), 449–460.

Smith, A. (2011). "Income inequality in the trust game." *Economics Letters*, 111(1), 54 – 56.

Snowdon, C. (2010). *The Spirit Level delusion: Fact-checking the left's new theory of everything*. Little Dice.

Steglich, C., Snijders, T. A., and Pearson, M. (2010). "Dynamic networks and behavior: Separating selection from influence." *Sociological Methodology*, 40(1), 329–393.

Steglich, C., Snijders, T. A., and West, P. (2006). "Applying SIENA: An illustrative analysis of the coevolution of adolescents' friendship networks, taste in music, and alcohol consumption." *Methodology: European Journal of Research Methods for the Behavioral and Social Sciences*, 2(1), 48–56.

Stevenson, B. and Wolfers, J. (2008). "Economic growth and subjective well-being: Reassessing the Easterlin paradox." *Brookings Papers on Economic Activity*, 39(1), 1–102.

Stiglitz, J. E., Sen, A., and Fitoussi, J.-P. (2009). "Report by the Commission on the Measurement of Economic Performance and Social Progress." (September).

Tanis, M. and Postmes, T. (2005). "A social identity approach to trust: Interpersonal perception, group membership and trusting behaviour." *European Journal of Social Psychology*, 35(3), 413–424.

The Economist (2009). "The power of mobile money." *The Economist*. Published on 24 September, 2009. Retrieved on 20 July, 2010, from http://www.economist.com/node/14505519.

The Economist (2010a). "Out of thin air." *The Economist*. Published on 10 June, 2010. Retrieved on 20 July, 2010, from http://www.economist.com/node/16319635.

The Economist (2010b). "The u-bend of life." *The Economist*. Published on 16 December 2010. Retrieved on 30 July, 2012, from http://www.economist.com/node/17722567.

The Economist (2014). "Bolivia's rentier republic." *The Economist, May 3rd 2014*.

Uganda Bureau of Statistics (2006a). "The 2002 Uganda population and housing census, economic characteristics." , Kampala: UBoS.

Uganda Bureau of Statistics (2006b). "The 2002 Uganda population and housing census, education and literacy." , Kampala: UBoS.

Uganda Bureau of Statistics (2006c). "The 2002 Uganda population and housing census, household characteristics." , Kampala: UBoS.

Uganda Bureau of Statistics (2006d). "The 2002 Uganda population and housing census, population size and distribution." , Kampala: UBoS.

Ugandan Communications Commission (2010). "Status of the communications market - March 2010." , Kampala: UCC.

Uslaner, E. (2008). *Corruption, Inequality, and the Rule of Law: The Bulging Pocket Makes the Easy Life*. Cambridge University Press.

Waverman, L., Meschi, M., and Fuss, M. (2005). "The impact of telecoms on economic growth in developing nations." *Policy Paper 2, 10-23*, Vodafone Group.

Wooldridge, J. M. (2002). *Econometric Analysis of Cross-Section and Panel Data*. MIT press.

Xiao, E. and Bicchieri, C. (2010). "When equality trumps reciprocity." *Journal of Economic Psychology*, 31(3), 456–470.

The Tinbergen Institute is the Institute for Economic Research, which was founded in 1987 by the Faculties of Economics and Econometrics of the Erasmus University Rotterdam, University of Amsterdam and VU University Amsterdam. The Institute is named after the late Professor Jan Tinbergen, Dutch Nobel Prize laureate in economics in 1969. The Tinbergen Institute is located in Amsterdam and Rotterdam. The following books recently appeared in the Tinbergen Institute Research Series:

546. M. SCHARTH, *Essays on Monte Carlo Methods for State Space Models*
547. J. ZENHORST, *Macroeconomic Perspectives on the Equity Premium Puzzle*
548. B. PELLOUX, *The Role of Emotions and Social Ties in Public On Good Games: Behavioral and Neuroeconomic Studies*
549. N. YANG, *Markov-Perfect Industry Dynamics: Theory, Computation, and Applications*
550. R.R. VAN VELDHUIZEN, *Essays in Experimental Economics*
551. X. ZHANG, *Modeling Time Variation in Systemic Risk*
552. H.R.A. KOSTER, *The internal structure of cities: the economics of agglomeration, amenities and accessibility*
553. S.P.T. GROOT, *Agglomeration, globalization and regional labor markets: micro evidence for the Netherlands*
554. J.L. MÖHLMANN, *Globalization and Productivity Micro-Evidence on Heterogeneous Firms, Workers and Products*
555. S.M. HOOGENDOORN, *Diversity and Team Performance: A Series of Field Experiments*
556. C.L. BEHRENS, *Product differentiation in aviation passenger markets: The impact of demand heterogeneity on competition*
557. G. SMRKOLJ, *Dynamic Models of Research and Development*
558. S. PEER, *The economics of trip scheduling, travel time variability and traffic information*
559. V. SPINU, *Nonadditive Beliefs: From Measurement to Extensions*

560. S.P. KASTORYANO, *Essays in Applied Dynamic Microeconometrics*
561. M. VAN DUIJN, *Location, choice, cultural heritage and house prices*
562. T. SALIMANS, *Essays in Likelihood-Based Computational Econometrics*
563. P. SUN, *Tail Risk of Equidity Returns*
564. C.G.J. KARSTEN, *The Law and Finance of M&A Contracts*
565. C. OZGEN, *Impacts of Immigration and Cultural Diversity on Innovation and Economic Growth*
566. R.S. SCHOLTE, *The interplay between early-life conditions, major events and health later in life*
567. B.N.KRAMER, *Why dont they take a card? Essays on the demand for micro health insurance*
568. M. KILIÇ, *Fundamental Insights in Power Futures Prices*
569. A.G.B. DE VRIES, *Venture Capital: Relations with the Economy and Intellectual Property*
570. E.M.F. VAN DEN BROEK, *Keeping up Appearances*
571. K.T. MOORE, *A Tale of Risk: Essays on Financial Extremes*
572. F.T. ZOUTMAN, *A Symphony of Redistributive Instruments*
573. M.J. GERRITSE, *Policy Competition and the Spatial Economy*
574. A. OPSCHOOR, *Understanding Financial Market Volatility*
575. R.R. VAN LOON, *Tourism and the Economic Valuation of Cultural Heritage*
576. I.L. LYUBIMOV, *Essays on Political Economy and Economic Development*
577. A.A.F. GERRITSEN, *Essays in Optimal Government Policy*
578. M.L. SCHOLTUS, *The Impact of High-Frequency Trading on Financial Markets*
579. E. RAVIV, *Forecasting Financial and Macroeconomic Variables: Shrinkage, Dimension reduction, and Aggregation*
580. J. TICHEM, *Altruism, Conformism, and Incentives in the Workplace*
581. E.S. HENDRIKS, *Essays in Law and Economics*
582. X. SHEN, *Essays on Empirical Asset Pricing*
583. L.T. GATAREK, *Econometric Contributions to Financial Trading, Hedging and Risk Measure-*

ment

584. X. LI, *Temporary Price Deviation, Limited Attention and Information Acquisition in the Stock Market*
585. Y. DAI, *Efficiency in Corporate Takeovers*
586. S.L. VAN DER STER, *Approximate feasibility in real-time scheduling: Speeding up in order to meet deadlines*
587. A. SELIM, *An Examination of Uncertainty from a Psychological and Economic Viewpoint*
588. B.Z. YUESHEN, *Frictions in Modern Financial Markets and the Implications for Market Quality*
589. D. VAN DOLDER, *Game Shows, Gambles, and Economic Behavior*
590. S.P. CEYHAN, *Essays on Bayesian Analysis of Time Varying Economic Patterns*
591. S. RENES, *Never the Single Measure*
592. D.L. IN 'T VELD, *Complex Systems in Financial Economics: Applications to Interbank and Stock Markets*
593. Y.YANG, *Laboratory Tests of Theories of Strategic Interaction*
594. M.P. WOJTOWICZ, *Pricing Credits Derivatives and Credit Securitization*
595. R.S. SAYAG, *Communication and Learning in Decision Making*