Freie Universität Berlin

Fachbereich Wirtschaftswissenschaft

Experimental and Survey Evidence on the Development of Preferences and Skills

Inaugural-Dissertation
zur Erlangung des akademischen Grades
einer Doktorin der Wirtschaftswissenschaft
des Fachbereichs Wirtschaftswissenschaft
der Freien Universität Berlin

vorgelegt von

Dipl.-Vw. Elisabeth Bügelmayer

aus

Hall in Tirol

Berlin, 2015

Erstgutachterin: Prof. Dr. C. Katharina Spieß

Zweitgutachter: Prof. Ronny Freier, Ph.D

Tag der Disputation: 23.05.2016

Erklärung zu Ko-Autoren

Die vorliegende Dissertation umfasst eine Einleitung (Kapitel 1) und drei Forschungspapiere (Kapitel 2 bis 4). Die Kapitel 1 und 4 wurden alleine verfasst. Die Kapitel 2 und 3 sind in Ko-Autorenschaft entstanden. Ko-Autorin des 2. Kapitels ist Prof. Dr. C. Katharina Spieß, Ko-Autor des 3. Kapitels ist Prof. Dr. Daniel Schnitzlein. Eine Liste mit Vorveröffentlichungen von Kapiteln befindet sich auf Seite 119.

Acknowledgments

First of all, I have to thank my main supervisor Prof. Dr. C. Katharina Spieß. I am extremely grateful to her for her continuous support and for her detailed feedback that helped to greatly improve my work. Whenever I needed her guidance, she took the time to listen and give advice. I benefited very much from her knowledge, experience and enthusiasm.

I am also very indebted to my second supervisor Prof. Ronny Freier for his great support. He encouraged me and always helped to improve my work with his expertise and very valuable comments.

Further, I especially thank the co-author of my second paper, Prof. Dr. Daniel Schnitzlein. I greatly enjoyed working with him and profited a lot from our collaboration.

Moreover, I want to thank my colleagues at the SOEP Department and the DIW Berlin for providing an inspiring research environment and giving me the opportunity to work with such a wealth of data. I also thank my friends and colleagues from the Graduate Center for the fruitful discussions and moral support.

Many friends from outside the DIW contributed to my work. Among all of them, I am especially grateful to Katharina Schultz for her optimism and help in the final stage of this dissertation.

Finally, I am fortunate to be part of a cheerful family whose various members supported me while writing this thesis. I am indebted to my mother for her unconditional support and for setting such a good example. Innumerable thanks to Benjamin, for being my companion through all phases of this thesis, and the kids for their patience and making me laugh.

Table of Contents

List of Tables

List of Figures

1 Introduction

"Field data, survey data, and experiments, both lab and field, as well as standard econometric methods can all improve the state of knowledge in the social sciences. There is no hierarchy among these methods and the issue of generalizability of results is universal to all of them."

(Falk & Heckman, 2009, p.537)

Understanding human decision-making is central to economics research and it is also the motivation behind writing this thesis. To learn more about why and how individuals make decisions, I analyzed survey, experimental and administrative data. The focus of my research is on decisions related to human capital formation and preferences crucial for social interaction – two topics that are very important in applied microeconomics. I analyze these decisions at different stages of the life-cycle, starting with other-regarding preferences in preschoolers, followed by skill formation in adolescents and preferences for honesty in adults. Answering the questions on skill formation and human capital accumulation was possible by exploiting the full potential the data provided, that is by combining different data types as well as making use of the data structure and analyzing paradata from the data collection process. Thus, this dissertation contributes to the literature by combining survey data, experimental data and paradata to analyze factors that influence individual decision-making. An important aspect this thesis focuses on are environmental factors that influence decision-making. These include family or community factors or situational factors in an experiment.

Economists, in their pursuit to understand and explain human behavior, have increasingly strived to identify causal relationships. The two main empirical tools applied for this intent are experiments and special econometric methods developed for causal evaluation. Experiments - carried out in the laboratory or in the field - allow for controlled variation, whereas causal econometric methods define identification assumptions to analyze interrelationships in observational data. Also, researchers have increasingly applied or combined these tools with survey data and registry data to gain empirical insights into human behavior. In line with these developments, this thesis draws on a combination of survey, experimental and administrative data. The advantages of each type of data are exploited to answer questions related to preference and skill formation at different stages of the life cycle.

Moreover, this dissertation adds to the literature by using the structure of the data as well as information generated in the data collection process to shed more light on preference heterogeneity and determinants of skill accumulation. In the following, I will shortly discuss different types of data and then describe how I analyze these types in my papers.

Laboratory experiments have experienced a sharp rise in popularity in economics in the last decades. Falk and Heckman (2009) report that before the mid 1960ies fewer than ten papers based on laboratory experiments were published. In 2008 the number of papers based on laboratory experiments rose to 4.15% of all published articles in three prominent economics journals. There are several advantages of laboratory experiments (see e.g,. Smith, 1994). Above all, the experimenter exerts a high degree of control over the decision making situation. The researcher determines payoffs, setting, subject pool and information sets. Furthermore, laboratory experiments allow for controlled variation in the parameter of interest which is a prerequisite for causal inference. Also, data from laboratory experiments can be replicated relatively easily compared to observational data.

In spite of their clear advantages, laboratory experiments have some shortcomings. Often, criticism concerning the generalizability of results is voiced as most laboratory experiments are carried out using rather homogeneous student subject pools. Another problem concerning the subject pool is the self-selection into experiments which can lead to biased results (Harrison et al., 2009). To complement the insights gained from laboratory experiments, *field experiments* have been increasingly used by economists (see Levitt & List, 2009). They provide a means to overcome some of the restrictions laboratory experiments face in predicting actual field behavior. To increase realism, nonstandard subject pools are recruited, field goods are used, and a field context is chosen (Harrison & List, 2004). Applying the taxonomy introduced by Harrison and List (2004), field experiments can be classified into artefactual, framed and natural field experiments. An artefactual field experiment is essentially the same as a laboratory experiment but is carried out with nonstandard subjects. Losing some degree of control but gaining more realism, a framed field experiment additionally includes elements of the naturally-occuring environment in the commodity, task, stakes or information sets of subjects. Finally, a natural field experiment is one where subjects naturally undertake experimental tasks and do not know that they take part in an experiment (Harrison & List, 2004). This also tackles the problem of self-selection into an experiment.

Another distinct data type is *survey data*. Survey data is collected with the purpose of drawing inference to a population of interest from a randomly selected sample representative of that population. Data can be collected via Internet surveys, by telephone, per mail or in person by an interviewer who visits the target person or household. The benefit of a survey is that the interviewer asks a standardized questionnaire on a variety of aspects of interest, such as the economic situation, health status or individual preferences and believes. However, the disadvantage of a survey is that actual behavior is not observed and can differ substantially from answers to the questionnaire (see e.g,. Dohmen et al., 2005 for the validation of a survey question with a field experiment). Also, individuals might not remember correctly when asked retrospective questions, or answers might suffer from social desirability (Fehr et al., 2002).

Recently, a strand of literature has emerged in economics that combines survey and experimental data to overcome some of the shortcomings of both data sources (see e.g., Dohmen et al., 2010; Bellemare & Kröger, 2007; Fehr et al., 2002). This combination of data is very promising for a number of reasons. First, it elicits behavior in incentivized experiments and allows linking the behavioral outcomes to a range of background variables. Thereby, the divergence between subjective information and actual behavior can be tested. Second, if experiments are embedded in long running panel data, relating behavioral outcomes to individuals' information over the life span may produce very interesting insights into decision-making. Third, carrying out experiments in representative surveys helps to reduce problems of the selectivity of the subject pool. A sample which is representative of the general population is certainly more heterogeneous than typical student subject pools. Of course, subjects can also refuse to participate in an experiment conducted in a survey. However, it is possible to rather precisely estimate the self-selection bias as various background variables of subjects who refuse to take part are available. Tanaka et al. (2010) use the term "broader studies" for studies that correlate experimentally elicited preferences with a large set of economic and demographic variables from a survey. A further advantage of these "broader studies" is, that the large number of results they produce helps to identify aspects which should be investigated more thoroughly by "targeted studies", i.e. studies that focus on one specific research question (Tanaka et al., 2010).

In contrast to survey data or experimental data, *registry and administrative data sources* have the advantage that they are very reliable and the sample often comprises a country´s whole population. Hartmann and Lengerer (2014) identify the following pros and cons compared to survey data. Registry and administrative data allow to break results down to small regional clusters, which is usually not possible with survey data. Furthermore, the data is collected continuously and unit non-response is very low. One of the disadvantages compared to surveys is that the data is rather inflexible and reacts slowly to societal change. Also, administrative or registry data are restricted to a few key figures and it is difficult to draw conclusions about peoples' personal opinions and preferences. Thus, combining survey data with administrative and registry data provides great potential for economists.

In addition, in every data collection process, data is generated which is not directly related to the research question, but can nevertheless provide interesting insights into respondents' behavior. This *paradata* has received increasing attention in the last decade as it can help to identify problems of specific survey questions, informs the researcher about non-response or sources of measurement error (Felderer et al., 2014). Examples of paradata are time-markers indicating how long respondents take to answer a question or key-stroke data. Another important type of paradata is information about interviewer characteristics, such as gender, age and education. In telephone or face-to-face surveys, where respondents interact with interviewers, interviewer characteristics and behavior can influence responses to a question (e.g., Schnell & Kreuter, 2005). Many German panel studies have recognized the potential of paradata to uncover unintended effects of a survey or to investigate interviewer effects and have started to collect it systematically (Felderer et al., 2014).

I now briefly introduce the three papers of my thesis ordered along the life-cycle before summarizing each in turn. In the second chapter, I (joint with my co-author) analyze an artefactual field experiment that was part of the German Socio-Economic Panel (SOEP) study and aimed at shedding light on the emergence of other-regarding preferences in a household context. The combination of experimental and survey data enabled us to relate child behavior in the experiment to extensive information about family background. In the third chapter, I (joint with my co-author) exploit the nature of the SOEP sampling process and a combination of survey and administrative data to investigate family and neighborhood influence on youth health and education. Finally, in the fourth chapter, I study an experiment on honesty preferences with adults embedded in the SOEP. To explain the displayed preferences for honesty, I draw on the paradata collected in the survey.

The *first paper*, "Spite and cognitive skills in preschoolers", investigates the emergence of other-regarding preferences and what factors are related to preference development. More specifically, we investigate the development of spiteful behavior in preschoolers. Spiteful preferences are important for the development of human cooperation (Jensen, 2010) and competitiveness (Balafoutas et al., 2012) and are very pronounced at the age of 5-6 years. However, little is known so far about the factors explaining spiteful behavior. We focus on investigating the link between cognitive ability and spite, as cognitive ability has been shown to be an important predictor for economic preference parameters (e.g., Dohmen et al.. 2010; Frederik, 2005). However, evidence on their relationship to other-regarding preferences is rather scarce. Furthermore, we are interested in the gender difference in the relationship between cognitive ability and spite, as gender differences in behavior have been shown to emerge early in life (Fehr et al., 2013).

The data we use stems from a series of pilot studies run to test a potential implementation of experiments in the SOEP. In this study, we analyze data from an experiment with five to six year old children carried out at their home. A sample of 214 mother-child dyads was visited at home by an interviewer. Mothers answered the standard SOEP household questionnaire as well as a personal and a child-specific questionnaire. In addition, their children took part in an experiment on other-regarding preferences and carried out tests of cognitive ability. This rare combination of survey and experimental data enabled us to thoroughly investigate the role of cognitive skills in the formation of spiteful preferences. Moreover, the sample of children we analyzed in their household context was supposedly more heterogeneous than laboratory samples are.

The experiment expanded the original design of Fehr et al. (2008) and consisted of four simple allocation decisions where each child had to decide how much suns (the experimental currency) to share with another (unknown) child. In all four decision tasks, children had the possibility to choose the equal distribution versus the following four alternatives: prosocial allocation, costly prosocial allocation, envy allocation and a costly envy allocation. In the costly prosocial and costly envy allocation children had to give up part of their own payoff to act prosocial or envious, respectively. Based on these four allocation decisions, we classified children into strongly spiteful and weakly spiteful types. In our sample, 15% of preschoolers were classified as strongly spiteful and 15% as weakly spiteful. The strongly spiteful type aimed at minimizing the other child's payoff whenever possible. Weakly spiteful types only minimized the other child`s payoff when it was not costly to themselves. There were no significant gender differences in spiteful types.

Additionally, two dimensions of children's cognitive ability were measured in two separate tests. First, children's crystalline intelligence was assessed with the Peabody Picture Vocabulary Test Revised (PPVT-R), a test on verbal skills. Second, fluid cognitive ability was measured by a subtest of the Culture Fair Intelligence Test (CFT1). This test assesses nonverbal reasoning competences and minimizes cultural influence. Also, we constructed a combined measure of cognitive ability. There were no gender differences in any of the three measures of cognitive ability we use.

In our regression analyses, we relate the measures of cognitive ability to the classification of spiteful types while controlling for a rich set of household, child and maternal characteristics. Our results show that children with higher cognitive skills are significantly more likely to act spitefully in the experiment. Analyzing the link between cognitive ability and spite more closely reveals that this association is mainly driven by fluid cognitive ability. This relationship is possibly explained by the fact that spiteful acts which reduce own as well as partner's payoff demand self-control and the capacity to calculate costs and benefits of this action. Also, there are obvious gender differences in the relationship between spite and cognitive ability. Both measures of cognitive skills are more strongly related to boys' spiteful behavior. The driving force behind this association might be a difference in preferences for competitiveness between boys and girls. These preferences have been shown to emerge early in life (Fehr et al., 2013) and might explain why boys with higher cognitive ability care more about their relative position to others.

In the *second paper*, we make use of a combination of SOEP data with administrative data to measure family and neighborhood influence on youth health and education. Next to assessing which of these two influences is more important, the combination of these data sources enables us to study pathways of neighborhood influence. These pathways have not been studied before as usually only individual survey data or aggregated administrative data is available.

Family and neighborhood factors influence children in a number of ways. How important these factors are for a child's success is an indication for the equality of opportunity in a society. If a child's life is to a large extent determined by its family and neighborhood background and to a lesser extent by own effort, then equality of opportunity is low. In the empirical economics literature, sibling correlations have been used to assess the equality of opportunity in a country. They are an expression for how much of the total outcome variance between siblings is due to shared family and neighborhood background. The higher the correlation is, the higher the influence of the shared background and the lower the equality of opportunity. As sibling correlations capture the importance of the shared background between siblings, they cover family and neighborhood influence. To disentangle these two factors, neighbor correlations have been used as they are a descriptive measure of how much of the variance in outcome between neighbors can be ascribed to shared neighborhood background.

In our study, we analyze youth education and health outcomes, as they are two crucial components of human capital formation. As measures of education outcomes we employ school grades and cognitive ability. Previous literature has found sibling correlations in educational background and cognitive ability to be between 0.40 and 0.61 in Norway, Sweden and the UK. In contrast, neighbor correlations in these outcomes ranged between 0.04 and 0.14 (Raaum et al., 2006; Lindahl, 2011; Nicoletti & Rabe, 2013). This implies that factors shared by siblings (including both family and neighborhood characteristics) are more important for explaining educational outcomes than shared neighborhood background. With the exception of Mazumder (2008, 2011) there is little evidence on sibling correlations in health outcomes. He finds that factors shared by siblings account for 27% to 40% of the variance in mental health and physical health outcomes of siblings in the US. Evidence on neighbor correlations in health outcomes is even less frequent. Thus, with our paper we aim at filling a research gap by estimating neighbor correlations in health outcomes.

In line with previous literature, our results show that neighbor correlations are substantially lower than sibling correlations. We estimate sibling correlations in school grades to be around 0.2 and neighbor correlations (in our preferred specification) around 0.04. Strikingly, we find neighbor correlations in cognitive ability around 0.2. Concerning health, sibling correlations in physical health and mental health outcomes range between 0.2 and 0.3 in our sample. Again, neighbor correlations are substantially lower and amount to 0.1 (mental health) and 0.06 (physical health). Our results show that neighborhood factors are more important in Germany than in other countries. We find this especially applies to adolescents' cognitive ability and mental health.

To uncover the driving factors behind the neighbor correlations, we merge administrative data on neighborhood characteristics on the postal code level to the SOEP. We find that neither the level of education nor the economic situation in a postal code area can explain the influence of the neighborhood.

To sum up, our paper contributes to the literature by providing sibling correlations and neighbor correlations in youth education and health outcomes for Germany. This enables a cross-country comparison of the relative importance of these factors. Also, we fill a research gap by estimating neighbor correlations in health outcomes. The relatively high neighbor correlation in mental health we find indicates that neighborhood influence on health outcomes is important but relatively under-researched so far. Lastly, with our combined survey and neighborhood data we undertake a novel approach to decompose the possible pathways of neighborhood influence.

In the *third paper*, I analyze data from an experiment on unobserved lying with adults implemented in the SOEP. In addition to combining survey and experimental data, in a novel approach I also draw on the paradata from the data collection process to explain behavior in the experiment. The experiment was designed to assess (aggregate) honest behavior in a heterogeneous sample of the German population. There are a number of papers studying honesty or cheating in the lab and in the field (e.g., Abeler et al., 2014; Fischbacher & Föllmi-Heusi, 2013; Houser et al., 2012). The rare combination of my data allows me to contribute to this literature by investigating the following research questions: Who lies? Who lies more/less when payoffs increase? And: Who do people lie to? By using paradata from the survey, I am able to shed more light on the importance of situational factors in the decision-making environment. Specifically, I focus on the characteristics of the interviewer and relationship to the interviewer to explain subjects' cheating in the experiment.

In the experiment, subjects rolled a two-coloured die and reported the outcome to the interviewer. The interviewer could not observe the outcome and thus did not know if subjects told the truth. Participants had an incentive to cheat because if the desirable colour of the die came up, they earned either €1 or €5, depending on which treatment they had been assigned to. As individual cheating is not observed, the actual reported proportion of the desirable outcome is compared to the probability of each outcome, i.e. 0.5, to infer subgroup specific cheating behavior. Overall, a non-negligible fraction of people cheated. In the total sample, 57% of subjects claimed to have rolled the desirable outcome, in the high stakes treatment even 62% reported the desirable outcome. Both proportions are significantly different from 50%. Only in the low stakes treatment the fraction of the desirable outcome is not statistically significantly different from the probability chance would predict. With respect to the first research question, descriptive results show that only individuals' education appeared to be related to cheating. Subjects with higher education were less likely to cheat. Next, to answer the second research question, I compare behavior in the high and the low stakes treatment. Whereas most subgroups cheated more in the high stakes treatment, my findings show that some individual characteristics, such as age and education, are associated with a robustness of cheating in the face of increased payoffs. Finally, I analyze who people lie to. Results show that subjects facing a female interviewer report the desirable outcome significantly more often than subjects interacting with a male interviewer. Also, older interviewers are more lied to. This result is confirmed by the regression analysis where I control for a number of background variables. Regressions also show that the longer the interview took, the less likely subjects were to cheat. These findings possibly uncover a form of discrimination against women and older people which has rarely been analyzed so far in this context. My results also highlight the importance of the decision-making environment for honest behavior. They indicate that it is crucial to take into account who people lie to when studying cheating behavior.

To conclude, in this thesis I use a combination of survey, experimental and administrative data sources to analyze preference and skill formation at different stages of the life-cycle. Combining these data sources is very promising for further research, as it helps to overcome shortcomings of each of these types of data. In addition, there are future research steps that may open new perspectives to better understand individual decision-making. One is, to embed experiments into longitudinal surveys. Although research has shown that integrating experiments into surveys is very promising, so far this is mostly done only for cross-sections investigations. Integrating experiments into longitudinal data would render possible very interesting insights into decision-making over the life-cycle. This allows decisions in an experiment to be analyzed with respect to incidences in subjects' lives that happened before or after the experiment. Also, the stability or malleability of preferences could be investigated more thoroughly by repeating experiments over the life-course. While the experiments I analyze in the first and third paper were carried out as part of the SOEP, longitudinal analyses are not possible. The sample of children was not integrated in the panel study and in the sample of the cheating experiment, individual cheating was not observed which allows only subgroup analyses but not analyses on the individual level.

In the second and third paper of this dissertation, the importance of context for decision-making is shown. From this perspective, it would be interesting to repeat the experiments in the laboratory and compare the outcomes in this controlled setting with subjects' behavior in their household context. Also, further research is needed to better understand the influence of neighborhood context on adolescents' outcomes, as results from the second paper show that in Germany this influence is not negligible. To this end, more data on smaller regional scales might help to shed more light on the role of the neighborhood for human capital formation.

1.1 References

Abeler, J., Becker, A., & Falk, A. (2014). Representative evidence on lying costs. *Journal of Public Economics*, *113*, 96–104. doi:10.1016/j.jpubeco.2014.01.005

Balafoutas, L., Kerschbamer, R., & Sutter, M. (2012). Distributional preferences and competitive behavior. *Journal of Economic Behavior & Organization, 83*(1), 125-135.

Bellemare, C., & Kröger, S. (2007). On representative social capital. *European Economic Review, 51*(1), 183-202.

Dohmen, T., Falk, A., Huffman, D., Sunde, U., Schupp, J., & Wagner, G. G. (2005). Individual risk attitudes: New evidence from a large, representative, experimentally-validated survey. *IZA Discussion Paper No. 1730.*

Dohmen, T., Falk, A., Huffman, D., & Sunde, U. (2010). Are Risk Aversion and Impatience Related to Cognitive Ability? *American Economic Review, 100*(3), 1238-60.

Falk, A., & Heckman, J. J. (2009). Lab experiments are a major source of knowledge in the social sciences. *Science, 326*(5952), 535-538.

Fehr, E., Bernhard, H., & Rockenbach, B. (2008). Egalitarianism in young children. *Nature, 454*(7208), 1079-1083.

Fehr, E., Fischbacher, U., von Rosenbladt, B., Schupp, J., & Wagner, G. G. (2002). A Nation-Wide Laboratory. *Schmollers Jahrbuch, 122*(519), 542.

Fehr, E., Glätzle-Rützler, D., & Sutter, M. (2013). The development of egalitarianism, altruism, spite and parochialism in childhood and adolescence. *European Economic Review, 64*, 369-383.

Felderer, B., Birg, A., & Kreuter, F. (2014). Paradaten. In: *Handbuch Methoden der empirischen Sozialforschung* (pp. 357-365). Springer Fachmedien Wiesbaden.

Fischbacher, U., & Föllmi-Heusi, F. (2013). Lies in disguise-An experimental study on cheating. *Journal of the European Economic Association, 11*(3), 525–547. doi:10.1111/jeea.12014

Frederick, S. (2005). Cognitive reflection and decision making. *Journal of Economic Perspectives*, 25-42.

Harrison, G. W., Lau, M. I., & Rutström, E. E. (2009). Risk attitudes, randomization to treatment, and self-selection into experiments. *Journal of Economic Behavior & Organization, 70*(3), 498-507.

Hartmann, P. H., & Lengerer, A. (2014). Verwaltungsdaten und Daten der amtlichen Statistik. In: *Handbuch Methoden der empirischen Sozialforschung* (pp. 907-914). Springer Fachmedien Wiesbaden.

Houser, D., Vetter, S., & Winter, J. (2012). Fairness and cheating. *European Economic Review, 56*(8), 1645–1655. doi:10.1016/j.euroecorev.2012.08.001

Jensen, K. (2010). Punishment and spite, the dark side of cooperation. *Philosophical Transactions of the Royal Society B: Biological Sciences, 365*(1553), 2635-2650.

Levitt, S. D., & List, J. A. (2009). Field experiments in economics: the past, the present, and the future. *European Economic Review, 53*(1), 1-18.

Lindahl, L. (2011). A comparison of family and neighborhood effects on grades, test scores, educational attainment and income—evidence from Sweden. *The Journal of Economic Inequality, 9*(2), 207-226.

Mazumder, B. (2008). Sibling similarities and economic inequality in the US. *Journal of Population Economics, 21*(3), 685-701.

Mazumder, B. (2011). Family and community influences on health and socioeconomic status: sibling correlations over the life course. *The BE journal of economic analysis & policy, 11*(3).

Nicoletti, C., & Rabe, B. (2013). Inequality in Pupils' Test Scores: How Much do Family, Sibling Type and Neighbourhood Matter? *Economica, 80*(318), 197-218.

Raaum, O., Salvanes, K. G., & Sørensen, E. Ø. (2006). The neighbourhood is not what it used to be*. *The Economic Journal, 116*(508), 200-222.

Schnell, R., & Kreuter, F. (2005). Separating interviewer and sampling-point effects. *Journal of Official Statistics*, 21, 3.

Smith, V. L. (1994). Economics in the Laboratory. *The Journal of Economic Perspectives*, 113-131.

Tanaka, T., Camerer, C. F., & Nguyen, Q. (2010). Risk and time preferences: linking experimental and household survey data from Vietnam. *The American Economic Review*, *100*(1), 557-571.

2 Spite and cognitive skills in preschoolers*

* This chapter is based on joint work with C. Katharina Spieß and has been published in the *Journal of Economic Psychology.* http://dx.doi.org/10.1016/j.joep.2014.10.001

3 Is it the family or the neighborhood? Evidence from sibling and neighbor correlations in youth education and health*

Abstract

In this paper we present sibling and neighbor correlations in school grades and cognitive skills as well as indicators of physical and mental health for a sample of German adolescents. In a first step, we estimate sibling correlations and find substantial influence of shared family and community background on all outcomes. To further disentangle the influence of family background and neighborhood, we estimate neighbor correlations. Our results show that for all outcomes, estimated neighbor correlations are clearly lower than estimated sibling correlations. However, especially for cognitive skills and mental health, neighbor correlations are still substantial in relation to sibling correlations. Thus, compared to existing results from other countries, the influence of the neighborhood is not negligible in Germany for these outcomes.

Acknowledgments:

We thank Karin Hederos Eriksson and participants of ESPE 2014, EALE 2014, and workshop participants at RWI Essen for helpful comments and suggestions. Elisabeth Bügelmayer gratefully acknowledges support from the Leibniz Association, Bonn, under the grant "Neighborhood effects: Analysis of individual rational behaviour in a social context."

* This chapter is based on joint work with Daniel D. Schnitzlein. A similar version has been published as SOEPpaper No. 716, http://dx.doi.org/10.2139/ssrn.2545766

We must not tolerate a situation where children cannot develop their talents because there is no equality of opportunity. We must not tolerate a situation where people have the impression that there is no longer any point in putting in any effort because they won't progress even if they work hard.

(Gauck, 2012)

3.1 Introduction

In his 2012 inauguration speech, German President Joachim Gauck formulated equality of opportunity as a normative policy goal to be reached. This normative goal, which is the foundation of most Western societies, not just Germany, implies that an individual's success is independent of factors beyond his or her control (Roemer, 1998).

The family and the neighborhood in which a child grows up are probably the two most important examples of such factors. Families influence their children in numerous ways, including the provision of resources, transmission of characteristics and skills, as well as investments in skill formation.[1] The neighborhood a child grows up in might influence child outcomes through various channels, such as social contagion and networks, environmental factors, as well as access to public services and infrastructure.[2]

In the economic literature the traditional approach to analyze equality of opportunity in a society is to estimate measures of intergenerational mobility – for example intergenerational correlations or intergenerational elasticities. This line of research has been extended to the analysis of sibling correlations (see for example Solon et al., 1991; Solon, 1999; Björklund & Jäntti, 2012). In contrast to intergenerational correlations – which only cover bivariate relationships – sibling correlations are a much broader measure of the influence of family and community background. In particular, sibling correlations measure the importance of all factors that siblings of one family share for their later outcomes. As this includes both, family background characteristics and neighborhood factors, sibling correlations can be seen as a measure of equality of opportunity as defined above.[3] The higher the similarity between siblings, the more important are family and neighborhood characteristics and the lower is the level of equality of opportunity.

---

[1] See for example Cunha and Heckman (2007, 2008), Currie (2009), Anger and Heineck (2010), Andrabi et al. (2012), Carneiro et al. (2013).

[2] See Galster (2012) and Durlauf (2004) for an overview of the literature on neighborhood effects.

[3] Note that sibling correlations measure only factors that are shared by siblings. Therefore, a sibling correlation is still a lower bound measure of the true influence of the family (Björklund and Jäntti, 2012). For more details, see section 3.3.

Authors estimate sibling correlations, among others, in monetary outcomes (e.g., Solon et al., 1991; Björklund et al., 2002; Mazumder, 2008; Schnitzlein, 2014), educational outcomes (Björklund & Salvanes, 2011; Nicoletti & Rabe, 2013), and health outcomes (Mazumder, 2011) for the US, UK, Germany and Scandinavian countries. The results show substantial influence of shared family and neighborhood background over all outcomes, but also significant cross-country differences in the importance of the combined effect.

The important question in interpreting these results is which factor is the main determinant? Or in other words, *is it the family or the neighborhood*? To address this question, analogous to sibling correlations, neighbor correlations are estimated. They provide a descriptive summary measure of how much of the outcome of neighbors can be attributed to the shared neighborhood. Solon et al. (2000), Page and Solon (2003), Raaum et al. (2006), Lindahl (2011) and Nicoletti and Rabe (2013) provide evidence for the US, Norway, Sweden and the UK, finding only weak neighbor correlations compared to sibling correlations for monetary and education outcomes.

In this paper, we focus on two crucial components of human capital formation, namely education and health. More specifically, we analyze school grades, cognitive skills, BMI, mental health, physical health, and height of a sample of German adolescents.

Regarding education, there are several studies investigating family and neighbor correlations. Solon et al. (2000) apply the method of sibling correlations to neighboring children. Their sample comprises 687 individuals from 379 families from the PSID. Sibling correlations in years of education are estimated to be 0.5. The correlation in educational attainment of neighboring children is 0.10 when controlling for basic family characteristics. Therefore, Solon et al. (2000) conclude that inequalities in educational attainment are mainly attributed to family background rather than growing up in the same neighborhood. Raaum et al. (2006) use census data from Norway to analyze the role of family and neighborhood on adult education and earnings. They report neighbor correlations in years of schooling of around 0.10. To adjust for the fact that similarities in adult outcomes might only be due to the fact that children growing up in the same neighborhood share similar family characteristics, again family background is controlled for. However, when adjusting for family background (parental education, family structure and parental income), neighbor correlations fall to around 0.04. Thus, childhood neighborhoods are substantially less important for adult outcomes than families. Lindahl (2011) assesses the importance of family versus neighborhood factors. She estimates correlations in income and education among siblings and children from the same neighborhood for about 13,000 individuals in Sweden. Applying a two-level model she finds that about 40% of the variation between siblings in years of education is due to shared background factors. Correlations among neighboring youth make up 15-20% of the sibling correlation. When family background is controlled for, neighborhood correlations drop to less than 3%. Lindahl (2011) concludes that family background is substantially more important than the neighborhood. This result is confirmed by Nicoletti and Rabe (2013) who also apply multilevel models to estimate the influence of family and neighborhood factors on pupil's test scores in the UK. At age 16, they report magnitudes of sibling and neighbor correlations of 0.61 and 0.14, respectively. These figures are comparable to the ones estimated by Solon et al. (2000) and by Raaum et al. (2006).

With respect to health, there is little evidence on sibling correlations. One exception is Mazumder (2011). He estimates sibling correlations in health using the PSID. Results show that correlations in siblings' outcomes are already high at birth and do not decline significantly over the life span. Even less research exists on neighbor correlations in health, although neighborhoods possibly affect health through physical neighborhood conditions such as pollution, poor quality housing, and stress of living in a dangerous neighborhood (see Robert, 1998). Further, the provision of medical infrastructure differs across neighborhoods.[4]

The empirical evidence for both education and health indicators in Germany is scarce. While Sieben et al. (2001) and Schnitzlein (2014) present sibling correlations in educational outcomes, neighbor correlations in education outcomes as well as sibling and neighbor correlations in health outcomes have not been analyzed for Germany so far. We focus on a sample of adolescents as they are often perceived as being most susceptible to disadvantageous neighborhood impacts (Kling et al., 2007). Therefore it is important to analyze the level of equality of opportunity this group faces. This, in turn, allows for a comparison of the extent of inequality in Germany versus other countries.

The contributions of our paper to the literature are threefold. First, we estimate sibling correlations in youth health and education and thus provide novel evidence on the joint

[4] Some evidence on the neighborhood impact on mental and physical health is provided by the Moving to Opportunity (MTO) study, a large randomized housing mobility experiment. Leventhal and Brooks-Gunn (2003) show that children benefitted most in terms of mental health when moving to a better neighborhood. Kling et al. (2007) find that teenage girls largely benefit in terms of mental health while teenage boys experience a decrease in these measures after moving to a lower poverty neighborhood. In the domain of physical health, they find no effect on adults (except obesity) and small effects on adolescents' physical health. Also using MTO data, Fortson and Sanbonmatsu (2010) focus on the influence of neighborhood quality on the physical health of children between 6 and 20 years of age. However, they do not find any positive impact on self-rated health, BMI, asthma or injuries in the medium-term.

importance of family background and neighborhood influence for youth outcomes in Germany. Second, we contrast the sibling correlations with estimated neighbor correlations in these outcomes. This enables us to give a first answer to the question of whether it is the family or the neighborhood that matters most in Germany. Third, we try to unveil the reasons and channels at work behind the neighborhood influence in Germany we find. To this end, we combine our survey data on adolescents with small scale neighborhood indicators from administrative register data. This data provides us with information about the average economic situation and the average level of education within a German postal code area and has rarely been employed so far.[5] We use this information to decompose the neighbor correlations with respect to neighborhood characteristics in order to investigate which pathways explain the neighborhood influence on our health and education outcomes.

We find substantial sibling correlations in youth education and health, albeit some of our estimated sibling correlations are lower than estimates for other countries. In line with the previous literature, our neighbor correlations are clearly lower than the estimated sibling correlations. However, our estimates show that for some youth outcomes, such as cognitive skills and mental health, the neighbor correlations still are substantial in relation to the sibling correlations. After controlling for parental characteristics, neither economic indicators nor the level of education in the neighborhood seem to be important determinants of these correlations.

The reminder of the paper is organized as follows. Section 3.2 discusses the empirical strategy. Section 3.3 discusses our data and our outcome variables. Section 3.4 contains our results and a discussion and section 3.5 concludes.

3.2 Econometric model and empirical strategy

Studies analyzing the neighborhood impact on children's outcomes face several econometric challenges due to non-random sorting of families into neighborhoods. Neighborhood characteristics and family characteristics are highly correlated as families self-select into neighborhoods. Due to this self-selection, simply regressing child outcomes on a set of family characteristics and neighborhood indicators makes it very difficult to disentangle the family

[5]Exceptions are Bauer et al. (2012), Bauer et al. (2013), and Hawranek and Schanne (2014).

effect from the neighborhood effect.[6] Solon et al. (2000) were the first to apply the method of sibling correlations to neighboring children as a solution to this problem: in an ideal world we would like to estimate the following model for the relationship between a child's (index i) outcome of interest y_{nfi} and his family and neighborhood characteristics as

$$y_{nfi} = \alpha' X_{nf} + \beta' Z_n + \varepsilon_{nfi} \quad (1)$$

with X_{nf} being a matrix of family (index f) variables, Z_n being a matrix of neighborhood variables (index n), α' and β' are the parameters to be estimated, with an error term, ε_{nfi}, that is uncorrelated with both family and neighborhood characteristics.

Based on the assumption that similar families tend to sort into similar neighborhoods (Tiebout, 1956), we expect X_{nf} and Z_n to be positively correlated. To obtain unbiased estimates of α and β from equation (1), ε_{nfi} has to be uncorrelated with both family and neighborhood factors. This only holds if we assume that X_{nf} and Z_n fully cover all relevant family and neighborhood information. As this is most certainly not the case, a simple estimation of equation (1) cannot answer the question raised in the introduction if family background or neighborhood factors are the most important determinant of a child's outcome.

We follow Solon et al. (2000) and adopt the idea raised in the literature on sibling correlations to estimate the importance of the neighborhood via the similarity of two individuals with identical community background. In particular, we estimate sibling and neighbor correlations in y_{nfi}. To compute these correlations, we first need an expression for the population variance of y_{nfi}. If we assume the model in equation (1), the population variance of y_{nfi} is the following

$$Var(y_{nfi}) = Var(\alpha' X_{nf}) + Var(\beta' Z_n) + 2Cov(\alpha' X_{nf}, \beta' Z_n) + Var(\varepsilon_{nfi}) \quad (2)$$

where $Var(\alpha' X_{nf})$ is the variance in family characteristics and $Var(\beta' Z_n)$ is the variance in neighborhood characteristics. These can be seen as the genuine family and neighborhood effects (Lindahl, 2011). $Cov(\alpha' X_{nf}, \beta' Z_n)$ instead measures the covariance between family

[6] In an overview of numerous studies applying this approach, Ginther et al. (2000) find a very large variation in results depending on the choice of the control variables and the selection of the neighborhood characteristics.

and neighborhood characteristics. Due to the nonrandom sorting, this covariance is assumed to be positive.

The covariance in y_{nfi} of two siblings, i and i', from the same family equals

$$Cov\left(y_{nfi}, y_{nfi'}\right) = Var\left(\alpha' X_{nf}\right) + Var(\beta' Z_n) + 2Cov\left(\alpha' X_{nf}, \beta' Z_n\right) \tag{3}$$

and the covariance in y_{nfi} of two neighboring children, i and i', from different families f and f' in the same neighborhood equals

$$Cov\left(y_{nfi}, y_{nf'i'}\right) = Cov\left(\alpha' X_{nf}, \alpha' X_{nf'}\right) + Var(\beta' Z_n) + 2Cov\left(\alpha' X_{nf}, \beta' Z_n\right). \tag{4}$$

Following the argumentation in Solon et al. (2000), we expect $Cov\left(\alpha' X_{nf}, \alpha' X_{nf'}\right)$ – the covariance between family characteristics in the same neighborhood – to be positive because of self-selection of families into neighborhoods. At the same time we expect $Var\left(\alpha' X_{nf}\right)$ to be larger than $Cov\left(\alpha' X_{nf}, \alpha' X_{nf'}\right)$ because children from the same family are more similar than children from different families. Therefore, we expect the neighbor correlation to be smaller than the sibling correlation:

$$\rho_{neighbor} = \frac{Cov\left(y_{nfi}, y_{nf'i'}\right)}{Var\left(y_{nfi}\right)} < \rho_{sibling} = \frac{Cov\left(y_{nfi}, y_{nfi'}\right)}{Var\left(y_{nfi}\right)} \tag{5}$$

Following Solon et al. (1991), we estimate the sibling and the neighbor correlation as intra-class correlations in a linear multilevel model. To illustrate this, suppose we observe an outcome y_{ig} for an individual i from group g. In this formulation, group can either be the family or the neighborhood. We now assume that y_{ig} can be modeled as

$$y_{ig} = \alpha_g + \mu_{ig} \tag{6}$$

where α_g is a group (family/neighborhood) specific component and μ_{ig} is an individual specific component. The group specific component covers all factors that are common to the individuals from the same group. In the case of siblings, these include not just family background factors, but also neighborhood background factors, and, in the case of neighbors, these include only neighborhood characteristics. With one individual being observed only in one group, these two components are orthogonal to each other. We therefore can write the variance of the outcome as sum of the variances of the two components:

$$Var(y) = Var(\alpha) + Var(\mu) \quad . \tag{7}$$

The correlation ρ between the outcome of two individuals, i and i', of one group then can be calculated as the ratio of the variance of the group specific component and the overall variance of the outcome:

$$\rho = \frac{Var(\alpha)}{Var(\alpha)+Var(\mu)} \tag{8}$$

Following Mazumder (2008), these variance components can be estimated via restricted maximum likelihood in the following model:

$$y_{ig} = \beta X_{ig} + \alpha_g + \mu_{ig} \tag{9}$$

where y_{ig} is the observed outcome of individual i in family/neighborhood g. X is a matrix of control variables containing year dummies, α_g is the family/neighborhood component, and μ_{ig} is the individual component. The standard errors of the sibling and neighborhood correlations are calculated via the bivariate delta method.

As can be seen from equation (4), the raw neighbor correlation covers the neighborhood effect as well as the effect of positive sorting of families into neighborhoods. Therefore a raw estimate of a neighborhood correlation would overestimate the importance of the neighborhood. To – at least partly – solve this problem, we present neighbor correlations adjusted for family background influences to account for the fact that similarities in neighboring adolescent's outcomes might only be due to similar family characteristics. We include family background variables in X_{ig} in equation (9) (Solon et al., 2000; Raaum et al., 2006; Lindahl, 2011). This – adjusted – neighbor correlation, which is still an upper bound estimate since we do not observe all relevant family characteristics, can now be compared to the sibling correlation to answer the question raised in the introduction.

In a next step, we add neighborhood characteristics to X_{ig}. This approach is suggested by Mazumder (2008) in order to uncover channels through which the influence of family background works. We adapt this approach for neighborhoods. If the added characteristics have an effect in determining the outcome y_{ig}, this will reduce the variance of the neighborhood component in equation (8), and thus reduce the neighbor correlation. The size of this reduction can be interpreted as a measure of the importance of this specific influence factor.

3.3 Data

Our study is based on German survey and administrative data. Our main data source is the German Socio-Economic Panel Study (SOEP). The SOEP provides detailed information on adolescents' health and educational attainment. It is an annual representative, longitudinal household panel with about 11,000 participating households and more than 20,000 individuals in the recent wave (Wagner et al., 2007). Since 2000, adolescents at age 17 fill in a youth specific questionnaire. Among other information, this questionnaire obtains detailed information about the youth's health and school performance.

3.3.1 Youth education and health outcomes

The SOEP data contains rich information on adolescents' educational performance. Specifically, as indicators of school performance, grades for both German and math of the last school report card are available. In Germany, grades range between 1 and 6, with 6 being the worst grade. In our estimations, we also include the average of these two grades as a combined measure of school performance.

Since 2006, the SOEP youth questionnaire includes a measure of cognitive ability. It consists of three parts, which were taken from the I-S-T 2000 (Solga et al., 2005), namely verbal skills, numerical skills, and abstract reasoning. The first two are measures of crystalline intelligence and the last measures fluid intelligence, i.e. innate ability (Anger, 2012; Richter et al., 2013). We use the sum of correct answers in all three tests as our measure of cognitive skills.

Self-reported height and weight, which we use to calculate the body mass index (BMI) for each individual, is also included in the youth questionnaire. As measures for physical and mental health we use the SF12v2 inventory, which, since 2002, is asked every second year in the SOEP. The SF12v2 contains twelve health questions, which are converted into continuous subscales of mental and physical health by a special algorithm (Andersen et al., 2007). As the SF12v2 questions are not available in the youth questionnaire, we use averages of the subsequent years in mental health status and physical health status in which the participants answer the regular SOEP questionnaire.

Furthermore, we include height as sole outcome variable in our analysis. Height should be mainly affected by genetic factors and therefore family background, but should be largely independent of the neighborhood (Duncan et al., 2005). It also serves as an indicator of comparableness of our results with existing studies. Table 3.1 shows descriptive statistics of our outcome variables.

3.3.2 Family and neighborhood characteristics

As discussed in the last section, raw neighborhood correlations would overestimate the influence of the neighborhood and, therefore, have to be adjusted by including family characteristics. Thus, we include the parental household income and the highest education achieved by either parent when the adolescents are 17 years old. Both indicators are available in the SOEP data. Table 3.2 shows descriptive statistics of these parental characteristics.

For a subset of our observations, we have detailed neighborhood information available. This information on the neighborhoods is mainly based on administrative data collected by the Institute for Employment Research (IAB) in Nuremberg. Specifically, our indicators, which are drawn from the official employment and unemployment registers, are available for the 2004 to 2010 period at the postal code level.[7] We merge our SOEP data with these administrative data on neighborhood characteristics. Available characteristics are the share of workers with a university degree in the workforce[8] and the share of workers with a

[7] The research data center of the Federal Employment Agency at the Institute for Employment Research provides the SIAB (vom Berge et al., 2013), which is a sample of these data and can be used either at the RDC or can also be obtained as a scientific use file. However, our neighborhood indicators are calculated from the full data, the Integrated Employment Biographies (IEB), not the sample.

[8] The workforce was computed as the sum of employed persons who are subject to social insurance contributions and registered unemployed persons (Bauer et al., 2011). Registered unemployed persons also include persons looking for work and persons who are liable to social security assistance if a family member is unemployed.

high school diploma in the workforce, which we use as indicators for the education level in the neighborhood. As indicators for the economic situation of the neighborhood, we include the share of unemployed in the workforce. In addition, we add the share of indebted people in the population over 18, which is provided by one of the largest private debt collection enterprises in Germany and is available for use with the SOEP data at DIW Berlin. Table 3.3 shows descriptive statistics of the neighborhood variables. Overall, they display substantial variation across postal code areas.[9]

3.4 Results

3.4.1 Sibling correlations in youth education and health

We begin the discussion of the results with the estimated sibling correlations. In Table 3.4, the estimated sibling correlations are shown alongside the estimated variance components as well as the number of families and observations. The sibling correlations in school grades in Germany are lower than expected, based on results from other countries. Our estimates of sibling correlations range from 0.17-0.23 for the German grade, the math grade, and the average of both. That means family and community background together explain between 17 and 23 percent of the variation in grades. Nicoletti and Rabe (2013) find sibling correlations in test scores of around 0.6 for the UK, while Mazumder (2011) estimates sibling correlations in high school GPAs of about 0.3 for the US. So, for our sample, the influence of family and community background on school performance is less pronounced than in the UK or the US.[10]

Further, we estimate a sibling correlation in cognitive ability of 0.46. So about half of the variation in cognitive ability can be attributed to factors shared by siblings. This estimate is comparable in size to studies using Swedish administrative data (Björklund et al., 2010; Björklund & Jäntti, 2012). Mazumder (2008) reports – based on data from the National Longitudinal Survey of Youth (NLSY) – only slightly higher correlations for measures of cognitive skills in the US.

People who are looking for work and are not registered as unemployed, self-employed and civil servants are not captured.

[9] Missing observations originate from data protection, which forbids reporting indicators for a postal code areas with fewer than 20 registered persons in the official administrative statistics.

[10] Note that grades are not directly comparable to test scores and might be a noisier measure than the test scores used in the previous literature.

The sibling correlations in our mental health and physical health outcomes range between 0.31 and 0.21, respectively. These results are lower than the 0.43 (mental health) and 0.37 (general health status) that Mazumder (2011) reports. However, health measures and age ranges are not directly comparable with Mazumder (2011). For BMI and height we estimate sibling correlations of 0.33 and 0.42. Concerning BMI, our estimates are comparable with findings from Mazumder (2008), who reports sibling correlations around 0.30. Regarding height, this is in line with the results Schnitzlein (2014) finds in an adult German sample and Duncan et al. (2001) find for adolescents in the US.

3.4.2 Neighbor correlations in youth education and health

Table 3.5 contains the corresponding neighbor correlations. The upper panel (Panel A) shows the raw *unadjusted* neighbor correlations and the lower panel (Panel B) shows neighbor correlations *adjusted* for family characteristics (parental education and household income) as discussed in section 3.2 to account for the possibility that outcomes of neighboring youth resemble each other not because of growing up in the same neighborhood but because they share the same family characteristics (see e.g., Raaum et al., 2006).

We find that the shared neighborhood accounts for around 8 percent of the variance in grades. Adjusting for parental background reduces the neighbor correlations to 0.06 - 0.08. So only between 6 and 8 percent of the variation in school grades can be attributed to the neighborhood.[11] In total numbers, this is even lower than the 0.14 Nicoletti and Rabe (2013) report for the test scores of 16 year olds in the UK. If we compare the sibling correlation to the neighbor correlation, the relation of both is roughly in the same range in the UK and in Germany.

Notably, neighbor correlations in cognitive ability are considerably higher and amount to 0.27. When family characteristics are controlled for, this correlation declines to 0.22. Still, about 22 percent of the variation in cognitive ability test scores can be attributed to the neighborhood.

Concerning health, the highest neighbor correlation we estimate is in mental health. Shared neighborhood factors at age 17 account for 14 percent of the variance in mental health. In contrast, neighbor correlations in physical health are around 0.09; adjusting for parental background leaves neighbor correlations in mental health and physical health largely

[11] Note that – as argued in section 3.3 – this is still an upper bound estimate of the influence of the neighborhood.

unaffected. Adjusted neighbor correlations in BMI amount to 0.10. Concerning height, we estimate an adjusted neighbor correlation of 0.15, which is considerably higher than the 1% Duncan et al. (2001) find.

This relatively high neighbor correlation in height seems suspect, unless one assumes sorting into neighborhoods based on height or neighborhood conditions that strongly affect inhabitants' height. We test the robustness of our results to rule out the possibility that our neighbor correlations are simply artifacts of our sibling correlations in neighborhoods with a small number of (large) families. To this end, we restrict our adjusted neighbor correlations to youth who live in neighborhoods in which we observe more than two families in the data. The results are presented in Table 3.6. Comparing these results to Panel B in Table 3.5 shows that neighbor correlations in height (and BMI) drop substantially, becoming insignificant. The other neighbor correlations slightly decline, except for cognitive ability. Thus, the substantial neighborhood influence on cognitive ability and mental health proves to be a robust result.

3.4.3 Comparison between sibling and neighbor correlations

Finally, we combine these results in Figure 3.1 to shed light on our question raised in the introduction. *Is it the family or the neighborhood?* The estimated sibling correlations, along with the adjusted neighbor correlations and the neighbor correlations restricted to youth living in neighborhoods with more than two families, are presented.

The first result is, as implied in equation (5), that the estimated neighbor correlations are all significant, but smaller in size compared to the estimated sibling correlations.[12] This result holds for all outcomes. If we compare the sibling correlations and the adjusted neighbor correlations, for all outcomes except cognitive ability and mental health, neighbor correlations are about one-third of the estimated sibling correlation. Adding the results from our robustness test shows a slightly increase in the difference between sibling and neighbor correlations. So the answer to the question is that it is *mostly* the family that influences the adolescent outcomes, with a *minor* part attributable to the neighborhood. Notably, for cognitive ability and mental health, the influence of the neighborhood is higher.

If we compare our results with existing results from other countries (Solon et al., 2000; Raaum et al., 2006; Lindahl, 2011; Nicoletti & Rabe, 2013) – although family influence is

[12] Technically, equation (5) refers to the unadjusted neighbor correlation, but the result also holds for the adjusted neighbor correlation.

most important – neighborhood influence is more important in relation to family influence in Germany. Thus, the results from the existing literature do not fully carry over to the German context.

In a last step, we aim at explaining channels through which the neighbor correlations we find might work. In a subsample of our youth data, we have detailed neighborhood characteristics available, including the level of education and the economic situation in the neighborhood. Following the decomposition approach by Mazumder (2008), we add these indicators to the estimation of the neighbor correlations. Further, we include community size as Page and Solon (2003) and Nicoletti and Rabe (2013) find significant effects of urbanicity on wages and test scores. This approach aims at disentangling the impact of shared family factors from shared community factors and is an extension of the previous literature, which often lacks joint information on family and neighborhood background.

The results are presented in Table 3.7 and show that adding these neighborhood characteristics does not decrease neighbor correlations after controlling for parental characteristics.[13] Thus, the educational structure and the economic situation in the neighborhood do not additionally contribute to the explanation of the influence of the neighborhood.

3.5 Conclusion

In this paper we present sibling and neighbor correlations in education and health for a sample of German 17 year olds. Moreover, we add to the literature by providing evidence on neighbor correlations in health outcomes, which, so far, have been scarcely investigated. Our analysis enables a cross-country comparison of the extent of inequality, which can be ascribed to growing up in the same family or in the same neighborhood in Germany.

Overall, we find evidence of substantial joint influence of family and community background on youth school grades, cognitive skills, mental health, physical health, BMI, and height. The estimated sibling correlations are partly lower than comparable estimates for other countries, but are, consistent with the literature, always larger than the neighbor correlations.

[13] Some of the neighbor correlations adjusted to parental characteristics are higher in the subsample than in our full sample. This is most likely due to the reduced sample as neighborhood characteristics are only available for the years 2004-2010. We estimated various specifications, including neighborhood variables sequentially or simultaneously but results remained nearly unchanged. Estimations are available upon request.

However, compared to existing results from other countries, the influence of the neighborhood is not negligible in Germany for some outcomes. This applies especially to the domains of cognitive ability and mental health, suggesting that, next to indicators of economic self-sufficiency, further research on the neighborhood impact to individual health is needed.

Our results have important implications for policy makers. While the existing literature emphasizes the importance of support for children with adverse family backgrounds, our findings suggest that neighborhood quality should also be a target of equality enhancing policies.

3.6 References

Andersen, H., Mühlbach, A., Nübling, M., Schupp, J., & Wagner, G. G. (2007). Computation of standard values for physical and mental health scale scores using the SOEP version of SF-12v2. *Schmollers Jahrbuch*, *127*.

Andrabi, T., Das, J., & Khwaja, A. I. (2012). What Did You Do All Day? *Journal of Human Resources*, *47*(4), 873–912. http://doi.org/10.3368/jhr.47.4.873

Anger, S. (2012). Intergenerational Transmission of Cognitive and Noncognitive Skills. In J. Ermisch, M. Jäntti, & T. M. Smeeding (Eds.), *From Parents to Children: The Intergenerational Transmission of Advantage* (pp. 393–421). New York: Russel Sage Foundation.

Anger, S., & Heineck, G. (2009). Do smart parents raise smart children? The intergenerational transmission of cognitive abilities. *Journal of Population Economics*, *23*(3), 1105–1132. http://doi.org/10.1007/s00148-009-0298-8

Bauer, T. K., Flake, R., & Sinning, M. G. (2013). Labor Market Effects of Immigration: Evidence from Neighborhood Data. *Review of International Economics*, *21*(2), 370–385. http://doi.org/10.1111/roie.12042

Bauer, T. K., Kasten, T., & Siemers, L. H. R. (2012). Business Taxation and Wages – Evidence from Individual Panel Data. *IZA Discussion Papers No 6717*. http://doi.org/10.2139/ssrn.2122520

Bauer, T. K., & Vorell, M. (2011). Neighborhood Effects and Individual Unemployment. *Ruhr Economic Papers*, (285).

Björklund, A., Eriksson, T., Österbacka, E., Jäntti, M., & Raaum, O. (2002). Brother correlations in earnings in Denmark, Finland, Norway and Sweden compared to the United States. *Journal of Population Economics*, *15*(4), 757–772. http://doi.org/10.1007/s001480100095

Björklund, A., Hederos Eriksson, K., & Jäntti, M. (2010). IQ and Family Background: Are Associations Strong or Weak? *The B.E. Journal of Economic Analysis & Policy*, *10*(1). http://doi.org/10.2202/1935-1682.2349

Björklund, A., & Jäntti, M. (2012). How important is family background for labor-economic outcomes? *Labour Economics*, *19*(4), 465–474. Retrieved from http://www.sciencedirect.com/science/article/pii/S0927537112000590

Carneiro, P., Meghir, C., & Parey, M. (2013). Maternal Education, Home Environments, and the Development of Children and Adolescents. *Journal of the European Economic Association*, *11*, 123–160. http://doi.org/10.1111/j.1542-4774.2012.01096.x

Cunha, F., & Heckman, J. J. (2007). The technology of skill formation. *American Economic Review*, *97*(2), 31–47.

Cunha, F., & Heckman, J. J. (2008). Formulating, Identifying and Estimating the Technology of Cognitive and Noncognitive Skill Formation. *Journal of Human Resources*, *43*(4), 738–782. http://doi.org/10.3368/jhr.43.4.738

Currie, J. (2009). Healthy, Wealthy, and Wise: Socioeconomic Status, Poor Health in Childhood, and Human Capital Development. *Journal of Economic Literature*, *47*(1), 87–122. http://doi.org/10.2307/27647135

Duncan, G., Boisjoly, J., & Mullan Harris, K. (2001). Sibling, peer, neighbor, and schoolmate correlations as indicators of the importance of context for adolescent development. *Demography*, *38*(3), 437–447. http://doi.org/10.1353/dem.2001.0026

Duncan, G. J., Kalil, A., Mayer, S. E., Tepper, R., & Payne, R. M. (2005). The Apple Does not Fall Far from the Tree. In S. Bowles, H. Gintis, & M. O. Groves (Eds.), *Unequal Chances: Family Background and Economic Success* (pp. 23–79). Princeton: Princeton University Press.

Durlauf, S. N. (2004). Neighborhood effects. In V. J. Henderson & J.-F. Thisse (Eds.), *Handbook of Regional and Urban Economics* (Vol. 4, pp. 2173–2242).

Fortson, J. G., & Sanbonmatsu, L. (2010). Child Health and Neighborhood Conditions: Results from a Randomized Housing Voucher Experiment. *J. Human Resources*, *45*(4), 840–864.

Galster, G. (2012). The mechanism (s) of neighbourhood effects: Theory, evidence, and policy implications. In M. van Ham, D. Manley, N. Bailey, L. Simpson, & D. Maclennan (Eds.), *Neighbourhood effects research: New perspectives* (pp. 23–56). Springer Netherlands.

Gauck, J. (2012). Inaugural Speech by Federal President Joachim Gauck on March 23, 2012 at the German Bundestag, Berlin.

Ginther, D., Haveman, R., & Wolfe, B. (2000). Neighborhood Attributes as Determinants of Children's Outcomes: How Robust Are the Relationships? *The Journal of Human Resources*, *35*(4), 603–642. http://doi.org/10.2307/146365

Hawranek, F., & Schanne, N. (2014). Your very private job agency: Job referrals based on residential location networks. Retrieved from http://www.econstor.eu/handle/10419/103076

Kling, J. R., Liebman, J. B., & Katz, L. F. (2007). Experimental Analysis of Neighborhood Effects. *Econometrica*, *75*(1), 83–119. http://doi.org/10.1111/j.1468-0262.2007.00733.x

Leventhal, T., & Brooks-Gunn, J. (2003). Moving to opportunity: an experimental study of neighborhood effects on mental health. *American Journal of Public Health*, *93*(9), 1576–82.

Lindahl, L. (2011). A comparison of family and neighborhood effects on grades, test scores, educational attainment and income—evidence from Sweden. *The Journal of Economic Inequality*, *9*(2), 207–226. http://doi.org/10.1007/s10888-010-9144-1

Mazumder, B. (2011). Family and Community Influences on Health and Socioeconomic Status: Sibling Correlations Over the Life Course. *The B.E. Journal of Economic Analysis & Policy*, *11*(3).

Nicoletti, C., & Rabe, B. (2013). Inequality in Pupils' Test Scores: How Much do Family, Sibling Type and Neighbourhood Matter? *Economica*, *80*(318), 197–218. http://doi.org/10.1111/ecca.12010

Page, M. E., & Solon, G. (2003). Correlations between Brothers and Neighboring Boys in Their Adult Earnings: The Importance of Being Urban. *Journal of Labor Economics*, *21*(4), 831–855. http://doi.org/10.1086/377021

Raaum, O., Salvanes, K. G., & Sørensen, E. Ø. (2006). The Neighbourhood is Not What it Used to be. *The Economic Journal*, *116*(508), 200–222. http://doi.org/10.1111/j.1468-0297.2006.01053.x

Richter, D., Metzing, M., Weinhardt, M., & Schupp, J. (2013). SOEP scales manual. *SOEP Survey Papers Series C - Data Documentations No 138*.

Robert, S. A. (1998). Community-Level Socioeconomic Status Effects on Adult Health. *Journal of Health and Social Behavior*, *39*(1), 18–37. http://doi.org/10.2307/2676387

Roemer, J. E. (2009). Equality of opportunity . Harvard University Press.

Schnitzlein, D. D. (2013). How important is the family? Evidence from sibling correlations in permanent earnings in the USA, Germany, and Denmark. *Journal of Population Economics*, *27*(1), 69–89. http://doi.org/10.1007/s00148-013-0468-6

Sieben, I. (2001). Family Background and Sibling Resemblance in Educational Attainment. Trends in the Former FRG, the Former GDR, and the Netherlands. *European Sociological Review*, *17*(4), 401–430. http://doi.org/10.1093/esr/17.4.401

Solga, H., Stern, E., Rosenbladt, B. von, Schupp, J., & Wagner, G. G. (2005). The

Measurement and Importance of General Reasoning Potentials in Schools and Labor Markets: Pre-Test Report. *DIW Research Notes*, *10*.

Solon, G. (1999). Intergenerational Mobility in the Labor Market. In O. Ashenfelter & D. Card (Eds.), *Handbook of Labor Economics* (Vol. 3, pp. 1761–1800). Amsterdam, New York and Oxford: Elsevier. http://doi.org/10.1016/S1573-4463(99)03010-2

Solon, G., Corcoran, M., Gordon, R., & Laren, D. (1991). A Longitudinal Analysis of Sibling Correlations in Economic Status . *Journal of Human Resources*, *26*(3), 383–392.

Solon, G., Page, M. E., & Duncan, G. J. (2000). Correlations between Neighboring Children in Their Subsequent Educational Attainment. *Review of Economics and Statistics*, *82*(3), 383–392. http://doi.org/10.1162/003465300558885

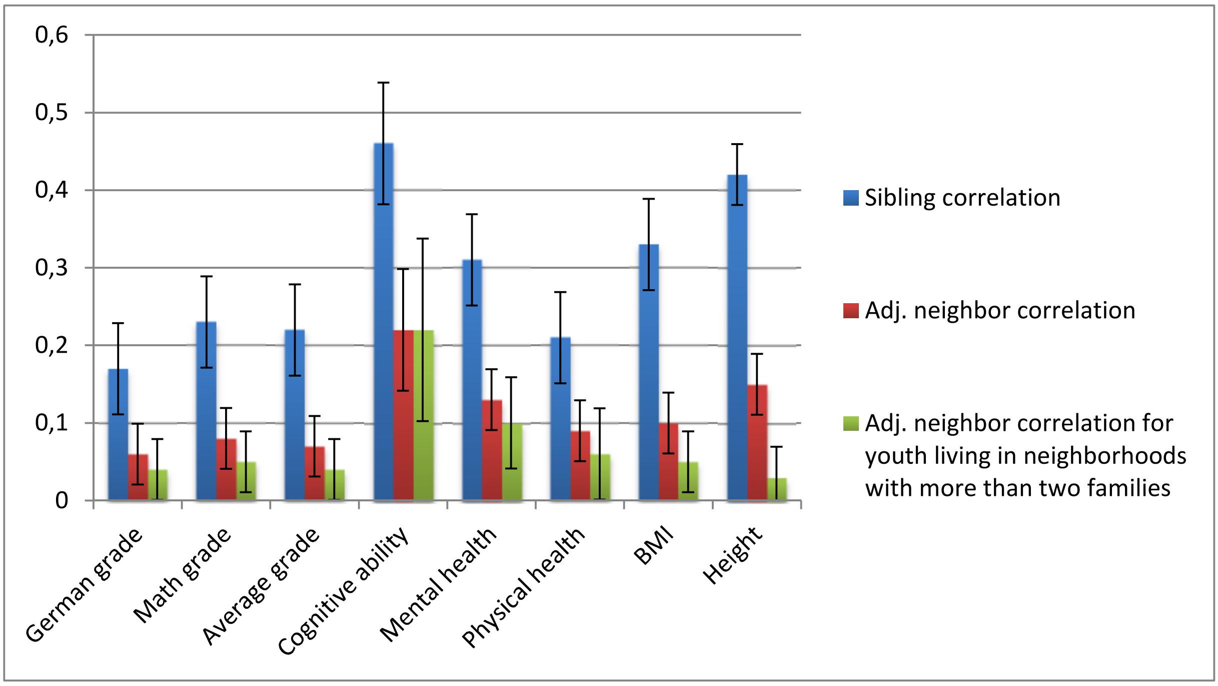
Tiebout, C. M. (1956). A Pure Theory of Local Expenditures . *Journal of Political Economy*, *64*(5), 416–424. Retrieved from http://www.jstor.org/stable/1826343?seq=1#page_scan_tab_contents

vom Berge, P., König, M., & Seth, S. (2013). *Sample of integrated labour market biographies (siab) 1975-2010*.

Wagner, G. G., Frick, J. R., & Schupp, J. (2007). The German Socio-Economic Panel Study (SOEP) - Scope, Evaluation and Enhancements. *Schmollers Jahrbuch - Journal of Applied Social Science Studies*, *127*(1), 139–169.

3.6.1 Figures and Tables

Figure 3.1 Sibling and neighbor correlations in youth education and health



Note: the figure shows estimated sibling and neighbor correlations in youth education and health; error bars indicate standard errors; results from REML estimation of a linear multilevel model; standard errors of sibling correlations and neighbor correlations are calculated via the delta method. Neighbor correlations are adjusted for family background characteristics. For details see Tables 3.4, 3.5 and 3.6.

Source: SOEP 2004-2010.

Table 3.1 Descriptive statistics of education and health outcomes

Variable	Mean	Minimum	Maximum	Standard deviation	Observations
German grade	2.89	1	6	0.82	3708
Math grade	2.97	1	6	1.03	3708
Average grade	2.93	1	6	0.76	3708
Cognitive ability	31.72	3	55	9.33	1715
Mental health	50.01	11.59	72.66	7.84	3088
Physical health	56.25	24.62	68.89	5.12	3088
BMI	21.81	13.05	47.88	3.43	3531
Height	173.21	125	204	9.13	3556

Source: SOEPv29 (2000-2012).

Table 3.2 Descriptive statistics of parental characteristics

Variable	Mean	Minimum	Maximum	Standard deviation	Observations
Household income	3102.47	10	35000	1756.48	3772
Highest parental education					
No school degree	0.01	0	1	0.12	3772
Secondary school degree	0.19	0	1	0.39	3772
Intermediate school degree	0.31	0	1	0.46	3772
Other degree	0.10	0	1	0.30	3772
University entrance diploma	0.07	0	1	0.26	3772
University graduate	0.31	0	1	0.46	3772

Source: SOEPv29 (2000-2012).

Table 3.3 Descriptive statistics of neighborhood characteristics

Variable	Mean	Minimum	Maximum	Standard deviation	Observations
Rate of unemployed %	11.35	2.25	36.21	6.23	1722
Rate of debtors %	9.99	2.40	34.18	3.64	1722
Rate of university graduates in the workforce %	7.83	1.41	38.57	4.61	1722
Rate of high school graduates in the workforce %	5.39	1.79	22.52	2.38	1722

Source: IEB 2004-2010; SOEPv29 (2004-2010).

Table 3.4 Sibling correlations in youth education and health outcomes

	German grade	Math grade	Average grade	Cognitive ability	Mental health	Physical health	BMI	Height
Sibling correlation	0.17 ***	0.23 ***	0.22 ***	0.46 ***	0.31 ***	0.21 **	0.33 ***	0.42 ***
	(0.03)	(0.03)	(0.03)	(0.04)	(0.03)	(0.03)	(0.03)	(0.02)
Family component	0.11	0.24	0.12	40.35	18.69	5.58	3.89	20.13
Individual component	0.54	0.82	0.44	46.55	41.54	20.63	7.76	27.68
Number of families	2664	2664	2664	1303	2249	2249	2524	2538
Number of individuals	3708	3708	3708	1715	3088	3088	3531	3556

Note: the table contains estimated sibling correlations in youth education and health; standard errors in parentheses; results from REML estimation of a linear multilevel model; standard errors of sibling correlations are calculated via the delta method.

Source: SOEPv29 2000-2012.

Table 3.5 Neighbor correlations in youth education and health outcomes

	German grade		Math grade		Average grade		Cognitive ability		Mental health		Physical health		BMI		Height	
A: Raw neighbor correlations																
Neighbor correlation	0.08	***	0.08	***	0.08	***	0.27	***	0.14	***	0.09	***	0.11	***	0.16	***
	(0.02)		(0.02)		(0.02)		(0.04)		(0.02)		(0.02)		(0.02)		(0.02)	
Neighborhood component	0.05		0.08		0.05		23.53		8.22		2.36		1.28		7.67	
Individual component	0.59		0.98		0.52		63.50		51.66		23.84		10.38		40.48	
Number of neighborhoods	1712		1712		1712		1039		1513		1513		1667		1673	
Number of individuals	3708		3708		3708		1715		3088		3088		3531		3556	
B: Neighbor correlations adjusted for household income and highest parental education																
Neighbor correlation	0.06	***	0.08	***	0.07	***	0.22	***	0.13	***	0.09	***	0.10	***	0.15	***
	(0.02)		(0.02)		(0.02)		(0.04)		(0.02)		(0.02)		(0.02)		(0.02)	
Neighborhood component	0.04		0.08		0.04		16.08		7.70		2.15		1.15		7.00	
Individual component	0.58		0.96		0.51		58.43		51.33		23.67		10.29		39.88	
Number of neighborhoods	1712		1712		1712		1039		1513		1513		1667		1673	
Number of individuals	3708		3708		3708		1715		3088		3088		3531		3556	

Note: the table contains estimated neighbor correlations in youth education and health; standard errors in parentheses; results from REML estimation of a linear multilevel model; standard errors of neighbor correlations are calculated via the delta method. Panel A contains raw neighbor correlations, panel B contains neighbor correlations adjusted for household income and highest parental education.

Source: SOEPv29 2000-2012.

Table 3.6 Adjusted neighbor correlations (sample restricted to youth living in neighborhoods with more than two families)

	German grade		Math grade		Average grade		Cognitive ability		Mental health		Physical health		BMI		Height
Neighbor correlation	0.04	*	0.05	**	0.04	**	0.22	***	0.10	***	0.06	*	0.05	**	0.03
	(0.02)		(0.02)		(0.02)		(0.06)		(0.03)		(0.03)		(0.02)		(0.02)
Neighborhood component	0.02		0.05		0.02		17.44		6.35		1.45		0.51		1.22
Individual component	0.60		1.00		0.54		63.30		55.68		23.94		10.23		40.85
Number of neighborhoods	267		267		267		215		262		262		266		266
Number of individuals	1240		1240		1240		524		1057		1057		1163		1174

Note: the table contains estimated neighbor correlations in youth education and health; standard errors in parentheses; results from REML estimation of a linear multilevel model; standard errors of neighbor correlations are calculated via the delta method. Neighbor correlations adjusted for household income and highest parental education.

Source: SOEPv29 2000-2012.

Table 3.7 Adjusted neighbor correlations (subsample with neighborhood characteristics)

	German grade		Math grade		Average grade		Cognitive ability		Mental health		Physical health		BMI		Height	
Panel A: Neighbor correlations adjusted for household income and highest parental education																
Neighbor correlation	0.09	**	0.09	***	0.09	**	0.21	***	0.19	***	0.09	**	0.13	***	0.14	***
	(0.04)		(0.03)		(0.03)		(0.05)		(0.04)		(0.04)		(0.04)		(0.04)	
Neighborhood component	0.06		0.10		0.05		15.88		12.07		2.27		1.48		6.50	
Individual component	0.57		0.94		0.50		59.19		50.42		23.12		10.17		40.43	
Number of neighborhoods	1026		1026		1026		818		934		934		1011		1014	
Number of individuals	1697		1697		1697		1263		1475		1475		1659		1672	

	German grade		Math grade		Average grade		Cognitive ability		Mental health		Physical health		BMI		Height	
Panel B: Neighbor correlations adjusted for household income,highest parental education, neighborhood level of education and economic situation																
Neighbor correlation	0.07	*	0.09	***	0.08	**	0.21	***	0.19	***	0.09	**	0.13	***	0.14	***
	(0.04)		(0.03)		(0.03)		(0.05)		(0.04)		(0.04)		(0.04)		(0.04)	
Neighborhood component	0.04		0.10		0.04		15.40		11.82		2.23		1.47		6.55	
Individual component	0.57		0.94		0.50		59.14		50.52		22.97		10.19		40.43	
Number of neighborhoods	1026		1026		1026		818		934		934		1011		1014	
Number of individuals	1697		1697		1697		1263		1475		1475		1659		1672	

Note: the table contains estimated neighbor correlations in youth education and health; standard errors in parentheses; results from REML estimation of a linear multilevel model; standard errors of neighbor correlations are calculated via the delta method. Panel A contains neighbor correlations adjusted for household income and highest parental education, panel B contains neighbor correlations adjusted for household income, highest parental education, unemployment rate, rate of debtors, rate of university graduates, rate of workers with high school diploma, dummy for living in a city with more than 500.000 inhabitants and a dummy for living in East Germany.

Source: SOEPv29 2004-2010; IEB 2004-2010.

4 Who lies and to whom? Experimental evidence on cheating in a household survey context

Abstract

In this paper, I evaluate a cheating experiment implemented in the German Socio-Economic Panel (SOEP) Study. My results show no relationship between cheating and socio-economic characteristics, personality traits, or economic preference parameters, except between cheating and education. I also study a high- and a low-stakes treatment to test the robustness of cheating in the face of increased gains. The results indicate that people cheat more when potential gains from doing so are greater, but that education appears to be a mediating factor. Further, I carry out a closer analysis of the decision-making situation focusing on the relationship between interviewer characteristics and the probability of cheating. This analysis provides the novel result that some interviewer characteristics—such as gender—significantly increase the probability of cheating.

Acknowledgments:

The cheating experiment was implemented in the SOEP in cooperation with Joachim Winter, LMU Munich. I am grateful for his contribution. I thank C. Katharina Spieß and Ronny Freier for very helpful comments.

4.1 Motivation

Honest behavior is a cornerstone of human social and economic interactions. Once the social norm of honesty has been internalized, most people adhere to this rule—even in situations where doing so is to their disadvantage. While this finding is intriguing for economists, who generally assume profit-maximizing behavior, it confirms what most of us observe in everyday situations: Most people do not ride the subway without buying a ticket, for instance, or take newspapers or food without paying for them, even when no one is watching (Levitt, 2006; Pruckner & Sausgruber, 2013).

Honesty and cheating[1] have been investigated in depth in a number of laboratory experiments (e.g., Fischbacher & Föllmi-Heusi, 2013; Houser et al., 2012; Mazar et al., 2008) and in the field (e.g., Abeler et al., 2014; Arbel et al., 2014; Ariely et al., 2014; Azar et al., 2013; Pruckner & Sausgruber, 2013). These studies document that a considerable portion of subjects face a moral cost of cheating and adhere to honesty when given the opportunity to cheat. Further, these studies demonstrate that the prevalence of cheating depends heavily on the subject pool used, the experimental design, and the decision-making environment.

Studying a cheating experiment implemented in a subsample of the German Socio-Economic Panel (SOEP) Study allows me to shed more light on these issues. I analyze a simple die-roll experiment with a high- and a low-stakes treatment where subjects self-report the outcome to the interviewer. As interviewers do not observe the true outcome of the die-roll, individual cheating cannot be detected and the extent of cheating in the sample is inferred from aggregate behavior.[2]

The combination of survey and experimental data has several advantages. First, my subject pool is more heterogeneous than typical student subject pools, as the sample I investigate is part of a representative household survey of the entire German population. Second, the SOEP study provides extensive information on socio-economic characteristics, personality, and preference parameters of decision-makers. Third, the SOEP also provides information on both the interview situation and interviewer characteristics. This offers me the opportunity to analyze a largely neglected aspect of the decision-making situation—namely, interviewer characteristics—and their role in the decision to cheat.

[1] In this paper, I use the terms “cheating” and “lying” interchangeably.

[2] This approach is also applied by, e.g., Houser et al., (2012) and Abeler et al., (2014). For more details, see Section 4.3.

Exploring a unique dataset, this paper contributes to the literature by shedding light on three main questions. First I ask: Who lies? That is, I analyze which characteristics of the decision-maker are related to cheating. Previous research in various experimental settings has indicated that cheating is related to personal characteristics of the decision-maker. For instance, Bucciol et al. (2013) reported that younger people and males are more likely to ride a bus without a ticket. Dreber and Johannesson (2008) showed that females are less prone to lie in a sender-receiver game. Further, religious affiliation appears to have a negative impact on the propensity to cheat among students in Israel (Arbel et al., 2014). Two studies on Germany found conflicting results. A telephone study by Abeler et al. (2014) found no significant cheating in a die-roll game, either in the sample as a whole or when looking at socio-demographic characteristics. In contrast, Ariely et al. (2014) showed that cheating is positively related to having an East German family background and age, whereas there is a negative association with education. In my sample, socio-demographic characteristics, personality, and economic preference parameters prove to be largely unrelated to cheating. The only exception is higher education, which lowers individuals' propensity to cheat.

Second, I am interested in learning more about the association between subjects' characteristics and cheating in the face of varying payoffs. A number of studies to date have investigated cheating behavior when stakes vary (e.g., Fischbacher & Föllmi-Heusi, 2013; Abeler et al., 2014; Gneezy et al., 2013; Gibson et al., 2013). While these studies analyzed how subjects in general respond to an increase in incentives, there is still a dearth of evidence regarding the relationship between specific decision-maker characteristics and varying stake size. Thus my second question is: Who cheats more/less when payoffs increase? My results show that cheating is more prevalent in the high-stakes treatment and that only the subgroup with higher education appears to be unaffected by an increase in incentives.

Third, this study provides novel insights into the importance of situational factors that have been neglected in the literature so far. I focus on interviewer characteristics to address the questions: Who do people lie to? Are some interviewer characteristics associated with a higher propensity of respondents to cheat? Previous research suggests that people's lying behavior is highly context-dependent. People lie less when the content of the lie is personal (Cappelen et al., 2013), when they have been treated unfairly in the past (Houser et al., 2012), or when others benefit from their cheating (Gino et al., 2013). However, little is known so far about who people lie to. This is surprising, as there are a few studies providing a first indication that characteristics of the cheater's counterpart such as gender play a crucial role in the decision to cheat. Rabinowitz et al. (1993), reported evidence from a field experiment and showed that shopkeepers were more likely to keep money that was apparently accidentally given to them by female customers than by male customers. Azar et al. (2013) also showed that gender plays a role when giving change, but found that customers acted more honestly towards female employees in a restaurant than male employees. In this study, I find that interviewer characteristics, in particular gender, age, and the relationship to the interviewer (measured by duration of the interview), are significant predictors of cheating behavior. For example, the probability that a respondent will report a desirable outcome is about 11% higher if he/she reports it to a female interviewer as compared to a male interviewer. This finding of higher dishonesty toward a female counterpart points to an aspect of discrimination against women that is still relatively unexplored.

This paper is organized as follows. In section 4.2, I give a brief overview of the related literature. Section 4.3 describes the data and experimental procedure. Thereafter, I present results in section 4.4 and discuss these results in section 4.5. In section 4.6, I summarize and conclude the paper.

4.2 Related literature

The literature on honesty, cheating or truth-telling spans the fields of economics, social psychology, and other areas of the social sciences (for a review, see Rosenbaum et al., 2014). In economics, experiments on honesty can be classified roughly into strategic and non-strategic experiments. Strategic experiments include sender-receiver games, where subjects can choose to communicate honest or dishonest messages to another subject. The gains of both players depend on whether the receiver believes the sender is honest and how the receiver acts in response. Non-strategic experiments include a wide range of settings, such as mail delivered to the wrong address that has to be forwarded to the right person or excess change given in shops or restaurants. The die-roll experiment I use is part of the popular subgroup of self-reported outcomes in the class of non-strategic experiments (e.g., Fischbacher & Föllmi-Heusi, 2013; Houser et al., 2012). Typically in this type of experiments, subjects report the outcome of a die roll or coin flip to the experimenter but only they know the true outcome. Cheating entails no cost to the decision-maker. The advantages of this approach are that participants' behavior is not confounded by strategic concerns.

Further, as subjects can conceal the real outcome from the experimenter, their actions are likely to be free from concerns of social desirability (Gneezy et al., 2013). The obvious disadvantage of this method is that individual cheating cannot be observed. The extent of cheating is computed by comparing the probability of each outcome, that is, 0.5 in the case of a coin-flip, with the actual rate reported.

I begin this section by giving a brief overview of the studies closest to mine in terms of experimental design. That is, I focus on studies with self-reported outcomes first in the lab and then in the field.

In the study by Fischbacher and Föllmi-Heusi (2013), subjects rolled a six-sided die and were paid according to the outcome. The outcome could not be observed by the experimenter. Comparing the reported number to the actual probability of a number to occur (i.e., 1/6), they found that 39% of subjects acted honestly and 20% cheated. Interestingly, a substantial fraction of subjects lied to some extent, that is, they cheated but did not report the highest outcome possible. Fischbacher and Föllmi-Heusi (2013) hypothesized that this behavior was driven by a desire to maintain a positive self-concept. Similarly, Mazar et al. (2008) found that participants did not cheat maximally (i.e., but rather engaged in incomplete cheating) when given the opportunity. In their set of experiments, participants had to solve simple search tasks in which they had to identify numbers in a matrix that added up to ten. In the cheating treatments, they self-reported their outcomes and thus had the opportunity to cheat. However, the level of cheating was far below the maximum extent possible. Results support the authors' theory that this behavior is due to people's wish to maintain a positive self-concept. Houser et al. (2012) investigated the relationship between fairness and cheating. After playing a dictator game, subjects had to toss a coin. The cheating rate of subjects who perceived themselves as being treated unfairly in the dictator game was 64% as opposed to 47% of those who rated their counterpart as fair. Their results show that the perception of being treated unfairly increases the individual's propensity to cheat.

In addition to the papers studying cheating in the lab, there has been an increase in recent years in studies investigating cheating in the field. The study most similar to mine in terms of sample and context is Abeler et al. (2014). They drew a random sample of the German population and conducted a cheating experiment over the telephone. In the telephone study, 658 subjects answered a questionnaire and took part in a coin-flip experiment at home. Abeler et al. (2014) did not find any significant over reporting of the desirable outcome. Also, they found no significant relationship between aggregate cheating and socio-demographic characteristics and economic preference parameters. In contrast, when replicating their study in the lab, they found considerably higher cheating rates. They assume that the reason behind this difference in behavior is that other norms are effective in a familiar home environment. Abeler et al. (2014) conclude that people face non-negligible lying costs that prevent them from cheating and that the extent of cheating depends on the decision-making environment. Further field evidence on cheating in the German population has been provided by Ariely et al. (2014). In a sample of 259 randomly selected individuals picking up their passports or ID cards from municipal offices, they found significant cheating in a die-roll experiment. Further, they showed that subjects cheated significantly more often if they had an East German (20%) family background as opposed to a West German background (10%). Also, age was positively correlated with cheating, while higher levels of education reduced the probability of reporting the desirable outcome. Ariely et al. (2014) ascribe the discrepancy in cheating behavior between East and West Germans to growing up in a socialist regime. Bucciol and Piovesan (2011) conducted a field experiment in an Italian summer camp with 182 children between the ages of 5 and 15. Overall, 77% of the children reported the desirable outcome in a simple die-roll experiment. None of the available socio-demographic characteristics were significantly related to the probability to cheat. Another field study on cheating was conducted by Azar et al. (2013). In a restaurant, they tested whether customers return excess change and find that two thirds of customers keep the excess change that the waiter returned apparently by accident. Returning the money was positively correlated with being a repeat customer and being a woman.

In order to investigate characteristics of cheaters more thoroughly, Bucciol et al. (2013) studied bus fare evasion in Italy. They collected socio-demographic characteristics from 541 randomly selected bus passengers and found that riding the bus without a ticket is positively correlated with being young, male, a non-European immigrant, unemployed, and relatively unconcerned with risk. In yet another study, Arbel et al. (2014) focused on the link between religiosity, gender, and cheating in Israel and found that cheating occurred most among secular females, whereas religious females displayed the highest levels of honesty.

Another aspect of my study is the variation in payoffs with the purpose of studying the robustness of cheating behavior among various subgroups. Evidence from the literature on subjects' propensity to cheat when incentivized with low and high payoffs is mixed. Abeler et al. (2014) implemented two treatments, one yielding €15 when reporting the desirable outcome and one yielding €20 when cheating maximally. However, in neither treatment did they find evidence of aggregate cheating or subgroup-specific cheating. Similarly, Fischbacher and Föllmi-Heusi (2013) tripled their payoffs but did not detect any differences in lying patterns compared to their baseline treatment. Among the studies reporting a negative association between incentives and cheating, Azar et al. (2013) varied the sum of change returned in their experiment in a restaurant and showed that subjects returned the lower amount of money less often. Mazar et al. (2008) also found that the level of honesty increased with the amount of payoffs. In contrast, Gneezy et al. (2013) found that lying is positively correlated with incentives in a considerable fraction of participants in a strategic experiment on lying aversion. This is in line with Freeman and Gelber (2010) and Gibson et al. (2013), who also showed that subjects respond positively to incentives to cheat.

Further, this study focuses on answering the question of who people lie to. That is, I analyze how interviewer characteristics are related to behavior in the cheating experiment. The economics literature dealing with experimenter effects is rather limited[3]. Some evidence of the effect of experimenter gender on subjects' behavior has been provided by Gneezy et al. (2009). They report a slightly higher willingness to compete if the experimenter was male in a non-Western society. With respect to experiments on honesty, two papers point to an interaction between subjects' gender and experimenter gender. Azar et al. (2013) found that females were more likely to return excessive change to female waitresses in a field experiment in a restaurant in Israel. In a field experiment in tourism shops in Austria, Rabinowitz et al. (1993) found that female shopkeepers were more prone to keep the change when overpaid by female customers. More research on the "gender-of-interviewer effect" has been pursued in the field of survey methodology. Groves and Fultz (1985) found that respondents tended to rate their economic outlook more positively when talking to male interviewers. Flores-Macias and Lawson (2008) also document gender-of-interviewer effects that appear to interact with respondents' gender.

[3] However, other fields such as pain research have long established a link between experimenter gender and behavior in experiments (see e.g., Levine & Lee De Simone, 1991).

4.3 Sample and experimental design

I analyze a short die-roll experiment with two treatments implemented in a subsample of the German Socio-Economic Panel (SOEP) Study. The SOEP is a representative panel of the German population with about 11000 participating households each year (Wagner et al., 2007). The past few years saw an increase in the implementation of experiments in the SOEP (e.g., Bügelmayer & Spiess, 2014; Kosse & Pfeiffer, 2012) which are all conducted by the same professional survey research institute that also carries out the main survey. The cheating experiment described below was carried out in 2010 in the second wave of a subsample of the SOEP. This subsample was randomly drawn from the German population with the purpose of testing new survey methods[4]. In 2010, it comprised a total of 1,175 households. In 96% of the households, one person agreed to participate in the experiment after completing the standard SOEP questionnaire.

The interview and experiment were set up as follows. The interviewer visited the SOEP household and every household member answered a detailed questionnaire. Interviews were carried out face to face with the help of a laptop in CAPI (Computer-Assisted Personal Interview) mode[5]. On average, interviews took 32 minutes per person. After the interviews, all respondents who had answered a personal questionnaire were eligible to take part in the cheating experiment. As only one participant per household was allowed, the person whose birthday was the most recent was selected.

The cheating experiment was introduced by the interviewer as a die-roll game with the purpose of finding out if there are some people who are luckier than others. Subjects were given a die with three black and three red sides and a cup. The interviewer handed the participant a short questionnaire (see Appendix B) and explained the rules of the game. Subjects were told to roll the die once in private and to not show the result to the interviewer. If red came up, subjects won the amount of euros indicated on the questionnaire. If black came up, they earned nothing. Neither was the possibility of cheating mentioned nor were subjects reminded to be honest. After rolling the die, subjects reported the outcome to the

[4] See Table 4.1 in Appendix A for a comparison of descriptive statistics of the subsample and the SOEP sample. While our sample is certainly more heterogeneous than typical laboratory subject pools, it still differs significantly from the core SOEP.

[5] Interviewers had to deviate from CAPI mode in 6% of the interviews. Those households were excluded from the analysis. Also, five individuals who participated in the experiment but provided no personal information were dropped. This results in the 1,088 observations used for my analysis.

interviewer who took the result down and immediately paid the participants according to their outcome.

Interviewers were trained carefully for the experiment. In the interviewer instructions, the experiment was termed a die-roll experiment and two reasons were given for conducting it. First, interviewers were told it was designed to test whether there are some groups of people who are luckier than others (the same explanation given to respondents). Second, interviewers were told that it serves as an additional incentive to take part in the overall survey as respondents would be given the chance to win a little extra money. Thus, interviewers did not know that it was an experiment on cheating. They were instructed not to influence the respondent in any way while rolling the die. Also, they were to make sure that they could not see the outcome of the die roll.

We ran a high- and a low-stakes treatment. Households were randomly selected into one of the two treatments by the survey research institute (Table 4.2 in Appendix A shows that there are no significant differences in socio-demographic characteristics of the two treatment groups). In the low-stakes treatment, subjects earned 1 euro if red came up. In the high-stakes treatment, the payoff was 5 euros if red came up. If black came up, participants in both treatments earned nothing. This variation in payoffs allows me to analyze the robustness of cheating behavior of different socio-economic subgroups in the face of increased gains.

The interviewer behavior as well as the cup and the die clearly signaled to participants that the actual die roll outcome could not be detected. Thus, cheating did not incur any costs to the participants and it could be concealed easily. As outcomes were not observed by the interviewer, individual cheating could not be detected. However, as the probability of each outcome is 0.5, aggregate outcomes inform us about the distribution of cheaters in our sample.

The combination of survey data with experimental outcomes has several advantages and makes an important contribution to the previous literature. The SOEP questionnaire covers a wide range of topics including employment, income, and health. It also includes personality and economic preference parameters, which equip me with an extensive set of background variables. Thus, cheating behavior in different subgroups of the sample can be investigated.

Further, previous studies hint at the malleability of cheating behavior in different contexts. Hence, I use the information on interviewers as well as the interview situation[6] that is provided by the SOEP to learn more about the characteristics of the interviewer that are associated with subjects' cheating. In addition to basic interviewer characteristics, such as gender or marital status that are assessed every year in the SOEP, in 2006 the SOEP carried out a detailed interviewer survey. In this survey, interviewers' socio-demographic characteristic as well as personality and economic preference parameters were assessed. To draw causal inference of the effect of interviewer's characteristics on subjects' cheating behavior, a crucial concern is the assignment practice of interviewers to households. In the SOEP, the assignment is effected first on the experience with other SOEP or longitudinal studies and second on the grounds of regional proximity. Other criteria, such as age or gender cannot be taken into account when first contacting a household because they are not known to the survey company in advance. After the first wave, priority is given to interviewer stability. That is, whenever possible the same interviewer should visit the household over the years.

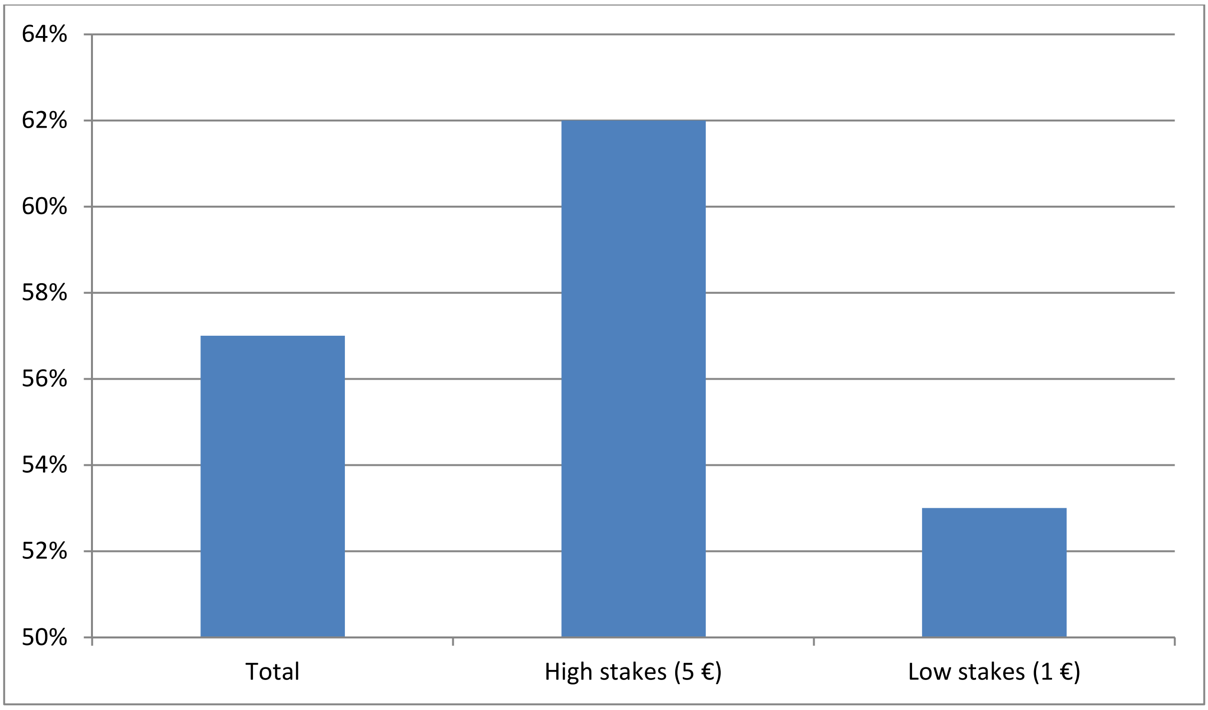
4.4 Results

4.4.1 Bivariate results

I begin by presenting results from the bivariate analysis of cheating behavior in the sample. With a fair two-coloured die and a sample of more than 1000 participants, the probability of each outcome is 0.5. However, in the sample 57% of subjects reported red (the desirable outcome). This proportion is significantly different from 0.5 (binomial test, p<0.000). Looking at the proportion of red reported by treatment reveals that subjects positively react to incentives to cheat. In the high-stakes treatment, 62% reported the desirable outcome, which is highly significantly different from 0.5 (binomial test, p<0.000). In contrast, in the low-stakes treatment the fraction of red reported is not statistically different from 0.5 (binomial test, p=0.19). Figure 4.1 summarizes these results.

[6] In particular, I analyze the duration of the interview. Time-markers of interviews and interviewer characteristics are examples of "paradata". This term stems from the field of survey methodology research and refers to data which is generated in the data collection process. Paradata has received increasing attention in the last years (see e.g., Felderer et al., 2014).

Figure 4.1 Die-roll outcomes – Percentage of red reported



Source: SOEP Innovation Sample 2010, own calculations.

Next, I compare aggregate cheating in different socio-economic subgroups. Table 4.1 presents the percentage of red reported along p-values from a binomial test and p-values from a two-sided test for equality of proportions. Several results emerge from Table 4.1. First, a significant fraction of people cheat when given the opportunity. The percentage of red reported is significantly different from 50% in nearly all social strata (p-values from binomial test). Exceptions are subgroups holding a university entrance diploma or a university degree, where the percentage of red reported is not distinguishable from a random die-roll. Second, personal characteristics are largely unrelated to aggregate cheating behavior. The percentage of red reported does not differ significantly between gender, high- and low-income, or within other subgroups (p-values from test for equality of proportions). Only subjects with higher education cheat significantly less than subjects without higher education, reflecting the first result.

Table 4.1 Proportion of red reported in different subgroups (total sample)

Subgroup		Proportion Red	N	p-value binomial test	p-value from test for equality of proportions
Gender	Female	0.592	613	0.000	0.138
	Male	0.547	475	0.043	
Age	> median	0.584	536	0.000	0.456
	<= median	0.562	552	0.004	
Income	> median	0.547	461	0.050	0.134
	<= median	0.595	479	0.000	
University degree	yes	0.504	244	0.949	0.018
	no	0.589	833	0.000	
University entrance diploma	yes	0.516	250	0.658	0.050
	no	0.586	826	0.000	
Married	yes	0.585	576	0.000	0.378
	no	0.559	512	0.009	
Region	East-Germany	0.605	223	0.002	0.267
	West-Germany	0.564	865	0.000	

Source: SOEP Innovation Sample 2010, own calculations.

Next, I analyze cheating separately by treatment. Behavior in the high-stakes treatment is similar to cheating in the total sample (Appendix A, Table 4.3). That is, subjects in all subgroups report the desirable outcome significantly more often than with probability 0.5. Again, the only exception is related to education. The percentage of red reported is not significantly different from 0.5 in the subgroup of subjects holding at least a university entrance diploma. As in the overall sample, slight differences within subgroups can be found with respect to the university entrance diploma.

In contrast, in the low-stakes treatment, cheating is not universal. In a number of subgroups, the desirable outcome is not reported significantly more often than chance would predict (Appendix A, Table 4.4). For example, males, younger people, and those with higher education display honest behavior in the aggregate. The only significant within-group difference can be found between individuals with and without a university degree.

In Table 4.2, I show how robust certain subgroups' cheating behavior was to increased incentives to cheat. In most subgroups, subjects cheated more in the high-stakes treatment.

However, there are certain socio-economic and socio-demographic characteristics that are associated with a robustness of cheating behavior in the face of increased gains. The cheating rate of subjects older than 50 did not differ significantly in the high- and low-stakes treatments (p=0.560). Similar to my earlier findings, education played a significant role with respect to subjects' cheating behavior. Subjects with higher education did not respond significantly to higher incentives to cheat (p=0.374).

Table 4.2 Comparison of red reported by subgroup and treatment – Socio-economic characteristics

Subgroup		Proportion Red Low stakes treatment	Proportion Red High stakes treatment	p-value from test for equality of proportions
Gender	Female	0.557	0.625	0.085
	Male	0.493	0.598	0.023
Age	> median	0.571	0.596	0.560
	<= median	0.491	0.631	0.001
Income	> median	0.507	0.585	0.092
	<= median	0.554	0.631	0.084
University degree	yes	0.474	0.531	0.374
	no	0.542	0.635	0.006
University entrance diploma	yes	0.442	0.577	0.035
	no	0.550	0.621	0.038
Married	yes	0.533	0.636	0.013
	no	0.525	0.589	0.145
Region	East-Germany	0.543	0.673	0.048
	West-Germany	0.526	0.599	0.029

Source: SOEP Innovation Sample 2010, own calculations.

I now turn to the relationship between cheating and interviewer-related characteristics (Table 4.3). In all subgroups, the proportion of red reported differs significantly from 50% (results from binomial test). Interestingly, the highest percentage of red reported occurs in the group of subjects with female interviewers. Female interviewers are lied to significantly more often than male interviewers. The difference in proportions between subgroups of male and female interviewers is nearly 10 percentage points and highly significant (test for equality of proportions, p=0.008). Also, the interviewer's age is related to aggregate cheating.

Interviewers who are older than the median are lied to slightly more (test for equality of proportions, p=0.055).

Table 4.3 Proportion of red reported by subgroup (total sample) – Interviewer characteristics

Subgroup		Proportion Red	N	p-value binomial test	p-value from test for equality of proportions
Interviewer gender	Female	0.622	405	0.000	0.008
	Male	0.536	563	0.092	
Interviewer age	> median	0.605	446	0.000	0.055
	<= median	0.544	522	0.049	
Interview duration	> median	0.549	446	0.042	0.212
	<= median	0.588	628	0.000	

Source: SOEP Innovation Sample 2010, own calculations.

4.4.2 Multivariate results

This section presents results from linear probability models on the relationship between cheating and individual characteristics as well as attributes of the decision-making environment. First, I regressed the probability of reporting red on individuals' socio-economic and sociodemographic characteristics to answer the question: Who lies? To shed more light on the second research question, I ran regressions separately by treatment. Finally, to answer the question of who people lie to, I added characteristics of the interviewer as they are potentially very important predictors of subjects' decision-making that have been neglected in the research to date. To test the robustness of results to model specification, I ran probit regressions for Model 1 and Model 2 (see Appendix A, Table 4.5). As coefficients from linear probability models and marginal effects from the probit regressions are very similar, in the rest of the paper I only present results from linear probability models.

The explanatory variables used here are gender, age, region, education, net household income and employment status. In the SOEP, personality is assessed by the Big Five (Lang et al., 2011). Preference parameters included in the SOEP comprise individuals' locus of control, risk aversion, as well as positive and negative reciprocity (see Richter et al., 2013 for a description of these concepts in the SOEP). Individuals' locus of control is related to various economic outcomes (see e.g., Cobb-Clark & Schurer, 2013 for an overview). Subjects with a higher external locus of control may be less likely to cheat because they perceive payoffs as "given" to a greater extent. Also, risk aversion has been associated with a number of economic outcomes (e.g., Belzil & Leonardi, 2007). In the context of cheating, Bucciol et al. (2013) found a positive relationship between risk aversion and riding a bus without a ticket. Further, as Fehr and Gächter (1998) stressed the importance of reciprocity for social norms, positive and negative reciprocity are included as explanatory variables.[7]

Model 1 in Table 4.3 depicts the relationship between the probability of reporting red—that is, the desirable outcome—and individual characteristics as well as personality and preference parameters. The results largely confirm findings from the bivariate analysis. Individuals' socio-economic characteristics, preference parameters, and personality traits are largely unrelated to reporting the desirable outcome in the experiment. Only subjects' agreeableness is positively and significantly related to the probability of reporting red.

I also ran estimations separately for the high- and the low stakes treatment to shed more light on the second research question. As in the total sample, most individual characteristics turned out not to be significantly related to cheating. Comparing results by treatment does not reveal substantial differences. Only agreeableness is positively related to cheating but just in the high-stakes treatment.

Finally, Model 2 investigates how interviewer characteristics and attributes of the interview situation are related to subjects' propensity to cheat. To this end, interviewer gender, interviewer age and interview duration were added to Model 1. Regression results confirm findings from the bivariate analysis. Specifically, they show that, in contrast to individual characteristics, situational factors are important predictors of cheating in the experiment. The probability of reporting red significantly increases when subjects face a female interviewer. Also, interviewer age is positively and highly significantly related to the probability of reporting red. Further, interview length has a significantly negative effect on cheating. The longer the interview takes, the less likely people are to report red.

[7] In a previous model I also included trust (see Richter et al. 2013 for a description of this concept in the SOEP) in the regressions. However, it was not significant and as it was asked only in 2012, the high number of missings minimizes the sample size drastically. Thus, I excluded it from further regressions.

Table 4.4 Probability of reporting red (LPM)

	(1) All	(1) High stakes treatment	(1) Low stakes treatment	(2) All	(2) High stakes treatment	(2) Low stakes treatment
Socio-economic characteristics						
Female	0.055	0.009	0.102	0.057	0.021	0.089
	(0.037)	(0.050)	(0.056)	(0.039)	(0.053)	(0.058)
Age	0.000	-0.002	0.002	-0.001	-0.003*	0.002
	(0.001)	(0.001)	(0.002)	(0.001)	(0.002)	(0.002)
University degree	-0.048	-0.082	-0.033	-0.043	-0.079	-0.004
	(0.046)	(0.062)	(0.069)	(0.048)	(0.065)	(0.071)
East	0.065	0.099	0.039	0.059	0.107	0.002
	(0.043)	(0.058)	(0.065)	(0.046)	(0.063)	(0.071)
Monthly net	-0.005	-0.003	-0.010	-0.003	-0.004	-0.004
household income	(0.013)	(0.018)	(0.020)	(0.015)	(0.020)	(0.023)
Unemployed	-0.030	-0.059	-0.007	-0.010	-0.085	0.073
	(0.066)	(0.086)	(0.098)	(0.069)	(0.095)	(0.101)
Married	0.050	0.037	0.052	0.052	0.056	0.046
	(0.038)	(0.052)	(0.058)	(0.041)	(0.055)	(0.063)
Preference parameters and personality						
Locus of control	0.002	0.017	0.002	-0.008	-0.003	-0.004
	(0.020)	(0.027)	(0.029)	(0.021)	(0.029)	(0.030)
Riskaversion	0.010	0.015	0.006	0.012	0.013	0.007
	(0.008)	(0.011)	(0.012)	(0.008)	(0.012)	(0.012)
Pos. Reciprocity	-0.021	-0.044	0.011	-0.007	-0.020	0.008
	(0.021)	(0.028)	(0.033)	(0.022)	(0.030)	(0.035)
Neg. Reciprocity	0.015	0.005	0.026	0.023	0.016	0.037
	(0.012)	(0.016)	(0.019)	(0.013)	(0.018)	(0.019)
Openness	0.003	-0.039	0.042	-0.002	-0.045*	0.040
	(0.015)	(0.020)	(0.022)	(0.016)	(0.020)	(0.024)
Conscientiousness	-0.001	0.028	-0.032	-0.001	0.022	-0.032
	(0.020)	(0.026)	(0.030)	(0.021)	(0.028)	(0.032)
Extraversion	0.022	0.054*	-0.015	0.017	0.054*	-0.022
	(0.016)	(0.021)	(0.025)	(0.017)	(0.023)	(0.026)
Agreeableness	0.047*	0.079**	0.019	0.051*	0.093**	0.012
	(0.019)	(0.026)	(0.029)	(0.021)	(0.029)	(0.031)
Neuroticism	0.024	0.042*	0.011	0.009	0.032	-0.001
	(0.015)	(0.019)	(0.021)	(0.016)	(0.022)	(0.022)

Table 4.4 Probability of reporting red (LPM) – (continued)

Decision-making environment						
Female interviewer				0.107**	0.070	0.177**
				(0.038)	(0.052)	(0.057)
Interviewer age				0.007***	0.008**	0.005
				(0.002)	(0.003)	(0.003)
Interview duration				-0.006***	-0.008**	-0.006*
				(0.002)	(0.003)	(0.003)
Constant	0.086	-0.045	0.095	-0.147	-0.315	0.043
	(0.240)	(0.302)	(0.390)	(0.277)	(0.356)	(0.442)
Observations	808	426	382	716	376	340
Adjusted R^2	0.009	0.038	-0.006	0.036	0.072	0.023
AIC	1168.658	600.748	572.993	1021.647	524.911	504.602
BIC	1248.466	669.673	640.065	1113.121	603.503	581.181

Notes: Estimated coefficients from Linear Probability Models, dependent variable equals one if subject reported red. Significance levels are denoted as follows: * p<0.10, **p<0.05, ***p<0.01, robust standard errors in parentheses. Source: SOEP Innovation Sample 2010, own calculations.

In the following, I attempt to shed more light on the relationship between cheating and interviewer characteristics. More specifically, I test the hypothesis that subjects behave differently towards interviewers that share similar characteristics to their own. Model 1 in Table 4.6 (Appendix A) investigates whether women are more likely to lie to other women. To this end, an interaction term between female respondent and female interviewer is added to check whether women behave differently towards female interviewers (in Table 4.6, I refrain from presenting results for each treatment; instead a treatment dummy is included). However, there is no evidence that the effect of interviewer gender on cheating differs between male and female respondents.

Further, I investigate whether subjects are more likely to cheat if the interviewer is similar in terms of age to the respondent. I include the age difference between interviewer and respondent as well as a dummy indicating if the interviewer is younger than the respondent in the regressions (Model 2 in Table 4.6, Appendix A). However, both these variables are insignificant.

As an additional robustness test, I include interviewers' personality in the regressions. However, the results remain largely unchanged when the interviewer's Big Five personality traits are added to regressions (Model 3, Table 4.6).

4.5 Discussion

Overall, I find that a non-negligible fraction of people dishonestly report the desirable outcome in a simple die-roll experiment carried out in their home. On the one hand, comparing the fraction of dishonest subjects in my study with a comparable lab experiment by Houser et al. (2012) shows that people in my sample cheat less. This discrepancy might be explained by a number of factors. First, my sample is very heterogeneous compared to the homogenous student subject pools usually analyzed in lab experiments. Second, the decision-making environment is inherently different in this study from the lab experiment. Subjects make decisions at home in a familiar environment and instantly report them to an interviewer. Third, it is very likely that they have established some kind of personal relationship to the interviewer as they answered a personal questionnaire taking a minimum of 30 minutes before the experiment.

On the other hand, cheating rates in my sample are higher than in Abeler et al. (2014), who do not find any evidence of cheating in their study with German households. Interestingly, Abeler et al. (2014) also study a heterogeneous subject pool at home and the length of interaction with the interviewer is comparable to my study. However, Abeler et al. (2014) conducted telephone interviews. Furthermore, for respondents to receive the incentives offered for their participation, they had to give their home address (or make note of an Amazon voucher code). My results might be an indication that the absence of cheating in their sample is due more to subjects' reluctance to give their home address than to their complete honesty.

Regarding the first research question—the relationship between individual characteristics and cheating—the only correlation I found in the bivariate analysis is between education and the probability of reporting the desirable outcome, red. This corresponds to the findings of Ariely et al. (2014) who also showed that more highly educated subjects cheat less. However, I cannot confirm the main conclusion of Ariely et al. (2014), namely that honesty is less prevalent among East Germans.

Regarding the second research question—the relationship between decision-maker characteristics and varying stake size—results from bivariate analysis show that older and more highly educated subgroups did not respond to higher incentives by cheating more. All other subgroups cheated significantly more when payoffs increased from €1 to €5. In the literature, there is a discussion on whether higher stakes lead to an increase in cheating because cheating has become more profitable or to a decrease in cheating because the higher stakes raise the moral costs. There is evidence for both hypotheses. My results are consistent with the findings of Gneezy et al. (2005) and Azar et al. (2013) and suggest that in the present sample, most people do not face increased moral costs of cheating (or that the increase in moral costs is not sufficient to stop people from cheating) when payoffs rise. However, as I analyze a between-subjects design and only observed aggregate cheating, I cannot fully answer this question. It might still be possible that some people in a within-subjects design do not cheat more when incentives are greater. Furthermore, it might be that the relationship between cheating and payoffs is hump-shaped, that is, that cheating increases up to some amount of payoff and then drops thereafter if payoffs are too high.

Further, as suggested by Houser et al. (2015), the high- and low-stakes treatment make it possible to compute the price elasticity of demand for honesty. In the context of honesty and cheating, the price elasticity of demand for honesty measures the percentage change of the quantity of honesty demanded when the price changes by one percent. That is, in this study, how elastic is the demand for honesty when the price increases from one to five euros? In the total sample, 53% reported the desirable outcome in the low-stakes treatment and 62% in the high stakes treatment. This corresponds to 6% and 24% cheaters (see Appendix C for more details about the calculation), respectively, and hence to 94% and 76% honest subjects. The percentage change of price from the low-stakes to the high-stakes treatment is -400%. Evaluating the price elasticity of demand for honesty at the point where p=1 (i.e., in the low-stakes treatment) of a linear demand curve yields a price elasticity of -0.048. This is very low, and in line with Houser et al. (2015), it suggests that the demand for honesty is rather inelastic.

I now turn to my results on the third research question: Who do people lie to? The most important variables explaining cheating in my study—even more important than individual characteristics and stake size—are dimensions of the decision-making environment (both in the bivariate and the multivariate analysis). The importance of these factors has been highlighted before (e.g., Cappelen et al., 2013; Abeler et al., 2014) but little is known about one crucial aspect of the choice situation: namely, the characteristics of the individuals who people choose to lie to. I find that in the aggregate, the desirable outcome is reported significantly more often to female interviewers and older interviewers. This is a novel finding, as the previous literature focused more on general discrimination against minorities (e.g., Ayres & Siegelman, 1995; List, 2004) or on gender differences in behavior of the decision-maker (e.g., Dreber & Johannesson, 2008).

One explanation for my results might be that a form of discrimination against women or older interviewers is at work in the interview setting. The works of List (2004) and Ayres and Siegelman (1995) refer to two theories of discrimination proposed by Becker (1971) and Phelps (1972), respectively: taste-based discrimination and statistical discrimination. In the context of the present study, taste-based discrimination would imply that female or older interviewers are lied to more as a means of compensating the costs respondents face when dealing with these groups (see Becker, 1971). Statistical discrimination would occur in the present context if respondents drew inferences from observable attributes of the interviewer, such as gender or age, regarding unobservable characteristics they believe the interviewer to possess (see Phelps, 1972). These beliefs about unobservable characteristics might be inaccurate and rooted in gender or age stereotypes. For example, respondents might think that women are more naive than men and therefore more likely to believe a lie. In an experimental study on deception in negotiations, Kray et al. (2014) show that women are deceived much more often than men. They suggest that women may be stereotyped as being more easily lied to and less capable of detecting a lie, possibly because they are perceived as “warmer” than men and therefore less willing to confront a liar, or as less competent. Similarly, older people face a number of negative stereotypes. They are regarded as less active, more old-fashioned, and more resigned (Gluth et al., 2010) or as slow in some contexts (Casper et al., 2011). Age-related stereotypes may also lead to the perception that older people are more easily lied to or more likely to believe a lie. However, I cannot test whether gender or age stereotypes drive differences in cheating as there is no information on how these groups are perceived. Thus, in the present study, I cannot disentangle the sources of discrimination.

Another possible explanation for the over-reporting of desirable outcomes to female interviewers could be that female interviewers were more likely to hint at the chance of cheating in the experiment. However, there are some facts that speak against this hypothesis. First, interviewers were very carefully trained in how to conduct the experiment including the exact wording they should use. Second, the longer the interview took, the lower the probability was of reporting the winning outcome. If female interviewers hinted at the possibility of cheating, than it is very likely that they did so to subjects with whom they had established a relationship. This in turn was probably the case with subjects they spent more time with, that is, subjects whose interviews took longer on average. However, the data shows the opposite, that is, cheating was lower in longer interviews. Third, as interviewers know a great deal about the financial situation of respondents, they might have been more likely to help individuals who are less well off. But neither household income nor local unemployment rates[8] have a significant effect on the probability of reporting the desirable outcome in the regressions. Taking all these arguments together, I am confident that subjects behave differently because of interviewer gender and not because of differences in behavior between male and female interviewers.

When looking more closely at experimenter-respondent gender pairs and cheating, two interesting facts emerge (Table 4.4). First, I find that female interviewers are lied to more by women than by men. Second, the fraction of red reported is not statistically different from 50% for male respondents with male interviewers. This implies that—in the aggregate—men do not lie to men.

Table 4.5 Percentage of red reported subject/interviewer gender

	Female interviewer	Male interviewer	p-value from test for equality of proportions
Female	65.04	57.65	0.077
Male	61.34	48.77	0.013

Source: SOEP Innovation Sample 2010, own calculations.

Similarly to our findings, Rabinowitz et al. (1993) show that female shop assistants were more likely to keep overpaid change from female tourists in an experiment in Austrian souvenir shops. In contrast, Azar et al. (2013) find that female customers more often return overpaid change to waitresses in a restaurant. It may be possible that women hold gender stereotypes regarding other women's naiveté or that they are more afraid of lying to men. Analyzing this can certainly yield interesting results but is not possible with my data.

Another important variable explaining cheating behavior in my analysis is interview length. I find that cheating significantly decreases with the duration of the interview. This hints at the influence of an interpersonal relationship on honesty. The longer subjects interact with the interviewer the more they might feel obliged to be honest. This hypothesis is

[8] Models which include local unemployment rates are not shown in the paper; results are available upon request.

supported by Chakravarty et al. (2011) who find that subjects lie less to friends than to strangers in a laboratory experiment.

While of limited statistical validity, yet another indication for the importance of the presence of the interviewer stems from a quick look at the cheating behavior of the 30 subjects initially excluded from the analysis. These individuals were excluded because there was no interviewer present when they filled out the questionnaire and rolled the die. Out of these 30 individuals, 83.33% reported the desirable outcome by mail.

4.6 Conclusion

There is a growing body of evidence from studies both in the lab and in the field showing that a considerable portion of people forego payoffs by acting honestly. However, the propensity to cheat appears to be strongly affected by the subject pool analyzed as well as the decision-making environment. The combination of survey and experimental data in the present study enables me to provide the following contributions to understanding individuals' cheating behavior. First, as my sample is part of the SOEP, extensive information on subjects' socio-economic characteristics is available, allowing me to compare the prevalence of cheating in different subgroups. Second, data on behavior in a high- and low-stakes treatment yields interesting insights into which subgroups of decision-makers are robust to cheating behavior in the face of increased gains. Finally, this paper adds to previous literature by exploring largely neglected features of the decision-making environment, in particular, characteristics of the person who people lie to.

In line with the literature, I find that a considerable fraction of subjects cheat in a simple die-roll experiment with self-reported outcomes where individual cheating cannot be detected. Looking at personal characteristics and the propensity to cheat, results document that they are largely unrelated. The only exception is education, which appears to reduce individuals' propensity to cheat and also serves as a mediating factor when stake size increases.

Comparing results from the high- and the low-stakes treatment clearly shows a positive correlation between payoffs and cheating in nearly all subgroups. This speaks against the hypothesis that people cheat less because of internal lying costs when payoffs increase. However, this possibility is not ruled out by the design of the present experiment. The cheating experiment I investigate has a between-subjects' design and only delivers results in the aggregate. Thus, there might be individuals who would cheat less in the high-stakes treatment but they would only be identifiable in a within-subjects design. Further, it might be that the relationship between cheating and incentives is hump-shaped. That is, cheating increases with incentives up to a certain threshold and decreases thereafter because internal lying costs grow too high. Thus, further research with a considerably greater variation in stakes is required.

In contrast to personal characteristics, aspects of the decision-making environment, in particular interviewer characteristics, are strongly related to cheating. Three interesting insights emerge from the analysis. First, female interviewers are lied to significantly more often. Second, the probability that subjects report the desirable outcome is positively correlated with the age of the interviewer. Third, the longer subjects and interviewer interact, the less likely subjects are to cheat.

Several conclusions can be drawn from these findings. First, when investigating cheating, it is crucial to keep in mind who subjects lie to. Personal characteristics of the experimenter or interviewer are potentially important determinants of behavior. Some experimenter characteristics—such as gender or age—can even represent a treatment of its own, and ignoring this may lead to incorrect conclusions.

Second, it might be that being lied to more often constitutes another form of discrimination faced by women and older people. However, this result remains to be validated in a more controlled setting. Further studies in the lab where the gender or age of the interviewer act as treatments have the potential of providing interesting insights. Also, it would be interesting to test the implications of this result in other settings, for example, professional or bargaining settings.

4.7 References

Abeler, J., Becker, A., & Falk, A. (2014). Representative evidence on lying costs. *Journal of Public Economics*, *113*, 96–104. doi:10.1016/j.jpubeco.2014.01.005

Arbel, Y., Bar-El, R., Siniver, E., & Tobol, Y. (2014). Roll a die and tell a lie – What affects honesty? *Journal of Economic Behavior & Organization*, *107*, 153–172. doi:10.1016/j.jebo.2014.08.009

Ariely, D., Garcia-Rada, X., Hornuf, L., & Mann, H. (2014). The (True) Legacy of Two Really Existing Economic Systems. *SSRN Electronic Journal*. doi:10.2139/ssrn.2457000

Ayres, I., & Siegelman, P. (1995). Race and Gender Discrimination in Bargaining for a New Car. *The American Economic Review*, *85*(3), 304–321. doi:10.2307/2118176

Azar, O. H., Yosef, S., & Bar-Eli, M. (2013). Do customers return excessive change in a restaurant? *Journal of Economic Behavior & Organization*, *93*, 219–226. doi:10.1016/j.jebo.2013.03.031

Becker, G. S. (1971). *The Economics of Discrimination*. *University of Chicago Press Economics Books*. University of Chicago Press.

Belzil, C., & Leonardi, M. (2007). Can risk aversion explain schooling attainments? Evidence from Italy. *Labour Economics*, *14*(6), 957-970.

Bucciol, A., Landini, F., & Piovesan, M. (2013). Unethical behavior in the field: Demographic characteristics and beliefs of the cheater. *Journal of Economic Behavior & Organization*, *93*, 248–257. doi:10.1016/j.jebo.2013.03.018

Bucciol, A., & Piovesan, M. (2011). Luck or cheating? A field experiment on honesty with children. *Journal of Economic Psychology*, *32*(1), 73–78. doi:10.1016/j.joep.2010.12.001

Bügelmayer, E., & Spiess, C. K. (2014). Spite and cognitive skills in preschoolers. *Journal of Economic Psychology*, *45*, 154–167. doi:10.1016/j.joep.2014.10.001

Cappelen, A. W., Sørensen, E. Ø., & Tungodden, B. (2013). When do we lie? *Journal of Economic Behavior & Organization*, *93*, 258–265. doi:10.1016/j.jebo.2013.03.037

Casper, C., Rothermund, K., & Wentura, D. (2011). The activation of specific facets of age stereotypes depends on individuating information. *Social Cognition, 29*(4), 393-414.

Chakravarty, S., Yongjin, M., & Maximiano, S. (2011). Lying and friendship. Unpublished.

Cobb-Clark, D. A., & Schurer, S. (2013). Two economists' musings on the stability of locus of control. *The Economic Journal*, 123(570), F358-F400.

Dreber, A., & Johannesson, M. (2008). Gender differences in deception. *Economics Letters*, *99*(1), 197–199. doi:10.1016/j.econlet.2007.06.027

Fehr, E., & Gächter, S. (1998). Reciprocity and economics: The economic implications of homo reciprocans. *European Economic Review*, *42*(3), 845-859.

Felderer, B., Birg, A., & Kreuter, F. (2014). Paradaten. In: *Handbuch Methoden der empirischen Sozialforschung* (pp. 357-365). Springer Fachmedien Wiesbaden.

Fischbacher, U., & Föllmi-Heusi, F. (2013). Lies in disguise-An experimental study on cheating. *Journal of the European Economic Association*, *11*(3), 525–547. doi:10.1111/jeea.12014

Flores-Macias, F., & Lawson, C. (2008). Effects of Interviewer Gender on Survey Responses: Findings from a Household Survey in Mexico. *International Journal of Public Opinion Research*, *20*(1), 100–110. doi:10.1093/ijpor/edn007

Freeman, R. B., & Gelber, A. M. (2010). Prize Structure and Information in Tournaments: Experimental Evidence. *American Economic Journal: Applied Economics*, *2*(1), 149–164. doi:10.1257/app.2.1.149

Gibson, R., Tanner, C., & Wagner, A. F. (2013). Preferences for Truthfulness: Heterogeneity among and within Individuals. *American Economic Review*, *103*(1), 532–548. doi:10.1257/aer.103.1.532

Gino, F., Ayal, S., & Ariely, D. (2013). Self-serving altruism? The lure of unethical actions that benefit others. *Journal of Economic Behavior & Organization,* 93, 285-292.

Gluth, S., Ebner, N. C., & Schmiedek, F. (2010). Attitudes toward younger and older adults: The German Aging Semantic Differential. *International Journal of Behavioral Development*, *34*(2), 147–158. doi:10.1177/0165025409350947

Gneezy, U., Leonard, K. L., & List, J. A. (2009). Gender Differences in Competition: Evidence From a Matrilineal and a Patriarchal Society. *Econometrica*, *77*(5), 1637–1664. doi:10.3982/ECTA6690

Gneezy, U., Rockenbach, B., & Serra-Garcia, M. (2013). Measuring lying aversion. *Journal of Economic Behavior & Organization*, *93*, 293–300. doi:10.1016/j.jebo.2013.03.025

Groves, R. M., & Fultz, N. H. (1985). Gender Effects among Telephone Interviewers in a Survey of Economic Attitudes. *Sociological Methods & Research*, *14*(1), 31–52. doi:10.1177/0049124185014001002

Houser, D., List, J. A., Piovesan, M., Samek, A. S., & Winter, J. (2015). On the Origins of Dishonesty: From Parents to Children. *NBER Working Paper Series* No. 20897

Houser, D., Vetter, S., & Winter, J. (2012). Fairness and cheating. *European Economic Review*, *56*(8), 1645–1655. doi:10.1016/j.euroecorev.2012.08.001

Kosse, F., & Pfeiffer, F. (2012). Impatience among preschool children and their mothers. *Economics Letters*, *115*(3), 493–495.

Kray, L. J., Kennedy, J. A., & Van Zant, A. B. (2014). Not competent enough to know the difference? Gender stereotypes about women's ease of being misled predict negotiator deception. *Organizational Behavior and Human Decision Processes*, *125*(2), 61–72. doi:10.1016/j.obhdp.2014.06.002

Lang, F. R., John, D., Lüdtke, O., Schupp, J., & Wagner, G. G. (2011). Short assessment of the Big Five: robust across survey methods except telephone interviewing. *Behavior Research Methods, 43*(2), 548-567.

Levine, F. M., & Lee De Simone, L. (1991). The effects of experimenter gender on pain report in male and female subjects. *Pain, 44*(1), 69–72. doi:10.1016/0304-3959(91)90149-R

Levitt, S. D. (2006). White-collar crime writ small: A case study of bagels, donuts, and the honor system. *American Economic Review*, *96*(2), 290–294.

List, J. A. (2004). The Nature and Extent of Discrimination in the Marketplace: Evidence from the Field. *The Quarterly Journal of Economics*, *119*(1), 49–89. doi:10.2307/25098677

Mazar, N., Amir, O., & Ariely, D. (2008). The Dishonesty of Honest People: A Theory of Self-Concept Maintenance. *Journal of Marketing Research, Vol. 45, No. 6, pp. 633-644, 2008*. doi:10.2139/ssrn.979648

Phelps, E. S. (1972). The Statistical Theory of Racism and Sexism. *American Economic Review*, *62*(4), 659–61.

Pruckner, G. J., & Sausgruber, R. (2013). Honesty on the Streets: A Field Study on Newspaper Purchasing. *Journal of the European Economic Association*, *11*(3), 661–679. doi:10.1111/jeea.12016

Rabinowitz, F. E., Colmar, C., Elgie, D., Hale, D., Niss, S., Sharp, B., & Sinclitico, J. (1993). Dishonesty, Indifference, or Carelessness in Souvenir Shop Transactions. *The Journal of Social Psychology*, *133*(1), 73–79. doi:10.1080/00224545.1993.9712120

Richter, D., Metzing, M., Weinhardt, M., & Schupp, J. (2013). SOEP scales manual (No. 138). *SOEP Survey Papers*.

Wagner, G. G., Frick, J. R., & Schupp, J. (2007). The German Socio-Economic Panel Study (SOEP) - Scope, Evaluation and Enhancements. *Schmollers Jahrbuch - Journal of Applied Social Science Studies*, *127*(1), 139–169.

Appendix A

Table A.4.1 Comparison of descriptive statistics SOEP Innovation Sample and SOEP core 2010

	Mean SOEP core	Mean Innovation Sample	Difference in means	N SOEP core	N Innovation Sample	t
Female	0.52	0.56	-0.04	14057	1088	-2.50*
Age	50.23	52.85	-2.62	14057	1088	-4.88***
City	0.47	0.28	0.19	18789	1088	13.53***
HH-Income	2860.35	2409.59	450.76	15478	940	8.25***
East	0.24	0.20	0.03	18783	1088	2.48*
College entrance diploma	0.21	0.23	-0.02	18272	1076	-1.44
University degree	0.23	0.23	0.00	13894	1077	0.25
Married	0.59	0.53	0.06	14057	1088	3.78***
Openness	4.42	4.63	-0.22	13091	982	-5.05***
Conscientiuosness	5.81	6.01	-0.20	13106	987	-6.52***
Extraversion	4.77	4.91	-0.14	13128	989	-3.62***
Agreeableness	5.33	5.48	-0.15	13158	988	-4.66***
Neuroticism	3.84	3.76	0.08	13174	990	2.01*

Source: SOEP core 2010, SOEP Innovation Sample 2010, own calculations.

Table A.4.2 Difference in means by treatment

	Mean Low stakes treatment	Mean High stakes treatment	Difference in means	N Low stakes treatment	N High stakes treatment	t
Female	0.57	0.56	0.00	527	561	0.13
Median age	52.79	52.91	-0.12	527	561	-0.12
Median HHincome	2396.94	2421.15	-24.20	449	491	-0.82
University degree	0.22	0.23	-0.02	522	555	-0.62
University entrance diploma	0.22	0.25	-0.03	522	554	-1.20
Married	0.54	0.52	0.02	527	561	0.73
East	0.22	0.19	0.03	527	561	1.20
Female interviewer	0.40	0.43	-0.03	469	499	-0.94
Interviewer median age	60.72	60.28	0.44	469	499	0.68
Median interview duration	31.84	32.34	-0.50	519	555	-0.82

Source: SOEP Innovation Sample 2010, own calculations.

Table A.4.3 Proportion of red reported in different subgroups in the high stakes treatment

Subgroup		Proportion Red	N	p-value binomial test	p-value from test for equality of proportions
Gender	Female	0.625	315	0.000	0.147
	Male	0.598	246	0.043	
Age	> median	0.596	282	0.002	0.191
	<= median	0.631	279	0.000	
Income	> median	0.585	236	0.011	0.290
	<= median	0.631	255	0.000	
University degree	yes	0.531	130	0.539	0.032
	no	0.635	425	0.000	
University entrance diploma	yes	0.577	137	0.087	0.355
	no	0.621	417	0.000	
Married	yes	0.636	291	0.000	0.255
	no	0.589	270	0.004	
Region	East-Germany	0.673	107	0.000	0.159
	West-Germany	0.599	454	0.000	

Source: SOEP Innovation Sample 2010, own calculations.

Table A.4.4 Proportion of red reported in different subgroups in the low stakes treatment

Subgroup		Proportion Red	N	p-value binomial test	p-value from test for equality of proportions
Gender	Female	0.557	298	0.056	0.147
	Male	0.493	229	0.895	
Age	> median	0.571	254	0.028	0.066
	<= median	0.491	273	0.809	
Income	> median	0.507	225	0.124	0.319
	<= median	0.554	224	0.000	
University degree	yes	0.474	114	0.640	0.199
	no	0.542	408	0.102	
University entrance diploma	yes	0.442	113	0.259	0.042
	no	0.550	409	0.048	
Married	yes	0.533	285	0.286	0.845
	no	0.525	242	0.480	
Region	East-Germany	0.543	116	0.403	0.738
	West-Germany	0.526	411	0.324	

Source: SOEP Innovation Sample 2010, own calculations.

Table A.4.5 Probability of reporting red (Probit)

	(1) All	(1) High stakes treatment	(1) Low stakes treatment	(2) All	(2) High stakes treatment	(2) Low stakes treatment
Socio-economic characteristics						
Female	0.054	0.007	0.102	0.056	0.020	0.088
	(0.036)	(0.049)	(0.053)	(0.038)	(0.051)	(0.055)
Age	0.000	-0.002	0.002	-0.001	-0.003*	0.002
	(0.001)	(0.001)	(0.002)	(0.001)	(0.002)	(0.002)
College degree	-0.047	-0.076	-0.033	-0.042	-0.075	-0.007
	(0.045)	(0.060)	(0.066)	(0.047)	(0.062)	(0.070)
East	0.066	0.105	0.037	0.060	0.117	-0.001
	(0.044)	(0.060)	(0.062)	(0.048)	(0.066)	(0.069)
Monthly net household income	-0.006	-0.004	-0.010	-0.003	-0.005	-0.004
	(0.013)	(0.016)	(0.020)	(0.013)	(0.017)	(0.022)
Unemployed	-0.030	-0.057	-0.005	-0.010	-0.082	0.073
	(0.066)	(0.090)	(0.098)	(0.069)	(0.094)	(0.101)
Married	0.051	0.041	0.052	0.053	0.056	0.048
	(0.038)	(0.051)	(0.057)	(0.040)	(0.054)	(0.060)
Preference parameters and personality						
Locus of control	0.002	0.017	0.002	-0.008	-0.005	-0.004
	(0.020)	(0.027)	(0.028)	(0.021)	(0.029)	(0.030)
Riskaversion	0.010	0.016	0.006	0.012	0.013	0.006
	(0.008)	(0.011)	(0.011)	(0.008)	(0.011)	(0.012)
Pos. Reciprocity	-0.021	-0.047	0.011	-0.008	-0.023	0.008
	(0.021)	(0.029)	(0.032)	(0.022)	(0.030)	(0.033)
Neg. Reciprocity	0.014	0.004	0.026	0.023	0.015	0.038*
	(0.013)	(0.017)	(0.019)	(0.013)	(0.018)	(0.019)
Openness	0.003	-0.041*	0.043	-0.001	-0.045*	0.040
	(0.015)	(0.020)	(0.022)	(0.016)	(0.021)	(0.024)
Conscientiuosness	-0.001	0.029	-0.032	-0.002	0.023	-0.033
	(0.020)	(0.026)	(0.030)	(0.021)	(0.027)	(0.031)
Extraversion	0.022	0.056**	-0.015	0.017	0.054*	-0.022
	(0.016)	(0.021)	(0.024)	(0.017)	(0.023)	(0.025)
Agreeableness	0.047*	0.080**	0.020	0.051**	0.092***	0.013
	(0.019)	(0.025)	(0.027)	(0.020)	(0.026)	(0.029)
Neuroticism	0.023	0.043*	0.011	0.009	0.031	-0.001
	(0.014)	(0.020)	(0.021)	(0.015)	(0.021)	(0.022)

Table A.4.5 Probability of reporting red (Probit) – (continued)

Decision-making environment						
Female interviewer				0.108**	0.070	0.177**
				(0.038)	(0.050)	(0.055)
Interviewer age				0.007***	0.008***	0.005
				(0.002)	(0.003)	(0.003)
Interview duration				-0.006***	-0.008**	-0.006*
				(0.002)	(0.003)	(0.003)
Observations	808	426	382	716	376	340
Pseudo R^2	0.02	0.06	0.03	0.05	0.09	0.06

Notes: Marginal effects from Probit regressions, dependent variable equals one if subject reported red. Significance levels are denoted as follows: * p<0.10, **p<0.05, ***p<0.01, standard errors in parentheses. Source: SOEP Innovation Sample 2010, own calculations.

Table A.4.6 Probability of reporting red (LPM) – additional models

	(1) Gender Interaction	(2) Age difference	(3) Interviewer personality
Respondent			
High stakes treatment	0.074*	0.077*	0.041
	(0.034)	(0.034)	(0.038)
Female	0.052	0.033	0.068
	(0.046)	(0.035)	(0.039)
Age	-0.001	0.000	-0.000
	(0.001)	(0.002)	(0.001)
University degree	-0.051	-0.051	-0.040
	(0.044)	(0.044)	(0.048)
East	0.037	0.036	0.042
	(0.044)	(0.044)	(0.049)
Monthly net household income	-0.012	-0.013	-0.000
	(0.013)	(0.013)	(0.015)
Unemployed	0.006	0.002	0.064
	(0.065)	(0.066)	(0.070)
Married	0.039	0.040	0.011
	(0.038)	(0.039)	(0.043)
Interaction between characteristics of respondent and interviewer			
Female # Female	-0.043		
	(0.070)		
Interviewer younger than respondent		-0.049	
		(0.061)	
Age difference to respondent		0.000	
		(0.002)	
Interviewer			
Female interviewer	0.140*	0.116**	0.116**
	(0.055)	(0.036)	(0.041)
Interviewer age	0.008***	0.007**	0.008***
	(0.002)	(0.002)	(0.002)
Interview duration	-0.005**	-0.005**	-0.008***
	(0.002)	(0.002)	(0.002)
Openness Interviewer			0.019
			(0.020)
Conscientiuosness Interviewer			-0.023
			(0.027)
Extraversion Interviewer			0.018
			(0.027)
Agreeableness Interviewer			0.038
			(0.024)
Neuroticism Interviewer			0.023
			(0.018)

Table A.4.6 Probability of reporting red (LPM) – additional models – (continued)

Constant	0.211	0.242	-0.148
	(0.135)	(0.143)	(0.324)
Observations	820	820	671

Notes: Estimated coefficients from Linear Probability Models, dependent variable equals one if subject reported red. Significance levels are denoted as follows: * p<0.10, **p<0.05, ***p<0.01, robust standard errors in parentheses. Source: SOEP Innovation Sample 2010, own calculations.

Appendix B

TNS Infratest Sozialforschung
Landsberger Str. 338
80687 München
Tel.: 089 / 5600 - 1009

Leben in Deutschland

Befragung 2010
zur sozialen Lage
der Haushalte

Würfelspiel

Wir möchten mit Ihnen nun ein kleines Würfelspiel durchführen. Dazu verwenden wir einen Würfel mit drei roten und drei schwarzen Seiten.

In unserem Spiel gilt: **„rot = Gewinn“** und **„schwarz = Niete“**.

Wie Sie sicherlich wissen, ist die Wahrscheinlichkeit „rot“ bzw. „schwarz“ zu würfeln gleich groß.

Manche Personen bezeichnen sich als Glückskinder, sagen also von sich selbst, dass sie oft Glück haben. Wir möchten nun untersuchen, ob es tatsächlich Personengruppen gibt, die häufiger Glück haben als andere.

Um das herauszufinden, möchten wir Sie bitten, mithilfe des Würfelbechers zu würfeln.
Falls Sie Glück haben, also **„rot“** würfeln, wird Ihnen ein Gewinn von 1 EURO ausbezahlt.

Um jegliche Form von Einflussnahme auszuschließen, soll die Interviewerin bzw. der Interviewer das Ergebnis nicht sehen. Bitte teilen Sie ihr/ihm lediglich das Ergebnis mit, also ob Sie „rot“ oder „schwarz“ gewürfelt haben.

☞ *Das Spiel wird nur mit einer Person pro Haushalt durchgeführt, die auch an der Befragung teilnimmt. Bei mehreren Teilnehmern an der Befragung soll diejenige Person würfeln, die zuletzt Geburtstag hatte. Das Würfelspiel findet nach der Beantwortung des Personen- bzw. Jugend-fragebogens statt.*

Bitte unbedingt eintragen lt. Adressenprotokoll:

Nr. des Haushalts: | | | | | | | |

Person Nr.: | | |

Vorname:

Bitte in Druckbuchstaben

Appendix C

In the low-stakes treatment, 53% report the desirable outcome and in the high-stakes treatment 64%, although chance would predict that 50% should roll the desirable outcome.

Low-stakes treatment: (53-50)/50=0.06, that is 6% cheat and 94 % are honest

High-stakes treatment: (62-50)/50=0.24, that is 24% cheat and 76 % are honest

Price elasticity of demand:

$$\mathrm{E_p} = \frac{\frac{\Delta \mathrm{Q}}{\mathrm{Q}}}{\frac{\Delta \mathrm{P}}{\mathrm{P}}} = \frac{\Delta \mathrm{Q}}{\Delta \mathrm{P}} \frac{\mathrm{P}}{\mathrm{Q}}$$

Price elasticity of demand for honesty at the point p=1 (i.e., low stakes treatment) of a linear demand curve:

$$E_p = \frac{\frac{94-76}{94}}{\frac{1-5}{1}} = -0.048$$

5 English summary

In this dissertation, I combine survey, experimental and administrative data to investigate preference and skill formation at different stages of the life-cycle. This rare combination of data helps to overcome shortcomings of each of these types of data. Surveys collect extensive information on individual characteristics, attitudes, health and many more. However, participants' actual behavior can deviate from their answers to a questionnaire. By embedding an incentivized experiment into a survey enables the researcher to elicit actual behavior and relate it to a number of background variables. Also, regional analysis may not be possible with survey data as usually the sample does not cover enough observations at a small regional scale. Enhancing survey data with data from administrative sources provides a way to relate individual characteristics to aggregated regional indicators.

In the *first paper*, I study with my co-author an experiment on other-regarding preferences in preschoolers. The experiment was part of a series of pilot-studies aiming at identifying the potential of integrating experiments in the SOEP. Thus, next to children's behavioral outcomes in four allocation decisions, we have a rich set of child and family background characteristics at hand. We focus on the development of spiteful preferences as they are especially prevalent at five to six years of age. Also, they have been shown to be important for human cooperation and competitiveness. We are especially interested in the relationship between cognitive skills and spitefulness, as cognitive ability has been found to be a predictor for economic preference parameters such as risk and patience. Further, we study if there is a gender difference in the relationship between cognitive ability and spiteful behavior as these differences emerge early in life. We find that both dimensions of cognitive ability, namely fluid and crystalline intelligence, are significantly related to spiteful behavior in the experiment on other-regarding preferences. In particular, higher fluid cognitive ability is associated with a higher probability of acting spitefully. This relationship is especially pronounced among boys. These findings suggest that spiteful behavior requires cognitive capacities and some amount of inhibitory control as in our experiment spiteful acts entailed a cost to the decision-making child. The results also show that especially boys with higher cognitive skills are prepared to reduce the other child's payoff in order to avoid being relatively worse off. One explanation for this behavior might be increased competitiveness among boys compared to girls. Differences in competitiveness between boys and girls have been found to emerge early in life and might be the reason for boys caring more about their position relative to another child.

In the *second paper*, we contribute to the literature on equality of opportunity by estimating the relative importance of siblings' shared family and neighborhood background on youth outcomes. We focus on health and education outcomes as they are crucial components of human capital formation. To shed more light on which factor – family or neighborhood background – better explains inequality in youth outcomes, we merge survey data on adolescents from the SOEP to administrative data from the official employment and unemployment registry. This provides us with a special dataset that allows us to relate youth outcomes to aggregate neighborhood characteristics at a small regional scale. In a first step we estimate sibling correlations in youth school grades, cognitive ability, physical health, mental health, BMI and height. Sibling correlations provide a measure for how much of the variance between siblings' outcomes can be accounted for by factors shared by them (i.e. family and neighborhood factors). A high correlation means that factors outside adolescents' control contribute more to their outcomes than individual factors and thus equality of opportunity is low. Sibling correlations in educational outcomes have been estimated to be between 0.4 and 0.6 for Norway, Sweden and the UK.

For Germany, we find that sibling correlations in education and cognitive ability are somewhat lower and range between 0.17 and 0.46. In a next step, we estimate neighbor correlations to analyze the importance of the shared community factors in comparison to family factors in more detail. Analogously to sibling correlations, neighbor correlations measure the importance of the shared neighborhood for youth outcomes. Previous studies have found neighbor correlations to be lower than sibling correlations, pointing at the fact that shared family background is more important than shared neighborhood background. For education outcomes, neighbor correlations have been shown to be between 0.02 and 0.14 in Norway, Sweden and the UK. We estimate neighbor correlations in education between 0.04 and 0.22. Neighbor correlations in health range between 0.05 and 0.10 in your sample. Thus, in line with previous literature, our estimated neighbor correlations are lower than the sibling correlations. Finally, we exploit the full potential of our dataset and aim at explaining the pathways of neighborhood influence. To this end, in a novel approach, we add neighborhood characteristics to decompose our estimations of neighbor correlations. However, we find that neither the economic situation nor the level of education at the post-code level can explain the neighbor correlations.

In sum, our paper contributes to the literature by estimating sibling as well as neighbor correlations for Germany. We provide evidence for the relative importance of family and neighborhood background for adolescents' outcomes. We find that it is the family that matters most. Also, we estimate neighbor correlations in health outcomes which have not been investigated so far and show that neighborhood factors play a role for youth mental health. Finally, the unique dataset we use enables us to attempt uncovering the channels at work behind the neighborhood influence we find.

The *third paper* studies preferences for honesty in a sample of adults in Germany. To this end, I analyze data from a cheating experiment carried out in a subsample of the SOEP. In the past years, economists have increasingly studied cheating because they are intrigued by the fact that rational individuals do not cheat when given the opportunity to and cheating cannot be detected. Similarly to the first paper, the combination of survey and experimental data allows me to relate behavior in the experiment to a number of personal characteristics of the decision-maker. However, in addition to the first paper, I also draw on the paradata from the data collection process to explain participants' behavior. This novel approach reveals very interesting insights into the role of the decision-making environment that have been largely neglected in the literature so far. More specifically, I focus on the personal characteristics and the relationship to the interviewer to explain cheating in the experiment. Participants were asked to roll a two-colored die and report the outcome to the interviewer. Depending on the treatment, they earned either €1 (low stakes) or €5 (high stakes) if the desirable color came up. Interviewers did not observe the true outcome of the die-roll and thus did not know if the participant behaved honestly. As individual cheating could not be detected, I infer the magnitude of dishonesty from aggregate behavior. That is, I compare the actual rate reported of the desirable outcome to 0.5, i.e. the probability of each outcome to occur in a fair two-colored die roll. Overall, 57% of participants reported the desirable outcome in the analysed sample. Comparing results by treatment reveals that participants positively react to incentives to cheat. That is, in the high stakes treatment 62% claimed to have rolled the desirable outcome versus 53% in the low stakes treatment. Further, I analyze if cheating is more prevalent in certain subgroups of decision-makers. I find that personal characteristics of participants are largely unrelated to cheating, with the exception of education which mediates dishonesty. In contrast, aspects of the decision-making environment, in particular interviewer characteristics, are strongly related to behavior in the experiment. More specifically, female interviewers and older interviewers are significantly more lied to. Also, the relationship to the interviewer appears to play a role. The longer participant and interviewer interact with each other; the lower is the probability of cheating. Several interesting conclusion emerge from these findings. First, increased dishonesty towards female and older people may point at a form of discrimination these groups face. Further, it is very important to take into account who people interact with in studies on cheating - an aspect which has been paid little attention to in the literature so far.

6 Deutsche Zusammenfassung

Diese Dissertation leistet einen Beitrag zur empirischen Mikroökonomie und untersucht die Entstehung von Präferenzen und Fähigkeiten in der frühen Kindheit, der Jugend und im Erwachsenenalter. Dazu wird das Potenzial von Survey-, Experimental- und administrativen Daten ausgeschöpft. Jede dieser Datenquellen hat ihre Vor- und Nachteile. Das Kombinieren verschiedener Datentypen ermöglicht es jedoch, die Stärken optimal für die Beantwortung meiner Fragestellungen zu nutzen. So werden in Umfragen umfassende Informationen zu den persönlichen Charakteristika, Meinungen, Gesundheit und vielen anderen Bereichen der Befragten erhoben. Ein Nachteil aber ist, dass tatsächliches Verhalten von den Angaben in einem Fragebogen abweichen kann. Daher ermöglicht das Einbetten von Experimentaldaten in eine Umfrage, tatsächliches Verhalten mit umfangreichen persönlichen Charakteristika in Verbindung zu setzen. Auch kleinräumige Regionalanalysen sind oft mit Umfragedaten nicht möglich, da die Fallzahl auf kleinen Ebenen nicht ausreicht. Die Anreicherung von Umfragedaten mit aggregierten administrativen Daten erlaubt es also, Individualdaten mit Regionalinformationen auf kleinräumiger Ebene in Beziehung zu setzen.

Die erste Studie trägt den Titel *„Spite and cognitive skills in preschoolers“*. Darin untersuchen wir die Entstehung von sozialen Präferenzen, im Besonderen von neidischem Verhalten, da dies im Vorschulalter besonders ausgeprägt ist und eine wichtige Rolle für zwischenmenschliche Kooperation spielt. Soziale Präferenzen bei Erwachsenen und Kindern wurden in der Literatur bereits untersucht, inwiefern diese aber mit kognitiven Fähigkeiten zusammenhängen wurde bisher jedoch vergleichsweise wenig beachtet. Ein Grund dafür mag das Fehlen von geeigneten Daten sein. Wir analysieren Daten aus einem Experiment mit fünf bis sechsjährigen Kindern, das im Rahmen einer Reihe von SOEP Pilot Studien durchgeführt wurde. Diesem Umstand ist es zu verdanken, dass zusätzlich zu den Experimentaldaten eine große Menge an Hintergrundinformationen zu den Kindern selbst, ihrer familiären Situation und dem Haushalt in dem sie leben, vorliegen. Diese seltene Kombination von Experimental- und Umfragedaten ermöglicht eine Analyse, die das Verhalten im Experiment mit sozio-demographischen und anderen Daten in Beziehung setzt. Zusätzlich zu ihrem Verhalten wurden in mehreren Tests die kognitiven Fähigkeiten der Kinder erhoben.

In dem Experiment zu sozialen Präferenzen wurden die Kinder in vier Allokationsentscheidungen aufgefordert, Sonnen (die Experimentalwährung) zwischen sich selbst und einem anderen Kind aufzuteilen. Je nachdem welche Aufteilung sie in den vier Allokationsentscheidungen wählten, wurden sie in verschiedene Präferenz-Typen eingeteilt. Die Typen die für die Analyse am wichtigsten waren, waren „sehr neidisch“ und „weniger neidisch“. Darüber hinaus wurden die Ergebnisse aus Tests zur fluiden und kristallinen Intelligenz der Kinder ausgewertet. In Regressionsanalysen wurde die Wahrscheinlichkeit zu einem der beiden neidischen Typen zu gehören mit den kognitiven Fähigkeiten sowie einer Reihe von Hintergrundvariablen in Verbindung gesetzt. Neben sozio-ökonomischen Variablen beinhalteten diese Kontrollvariablen unter anderem auch die Persönlichkeit und das Sozialverhalten der Kinder. Die Resultate zeigen, dass Kinder mit ausgeprägten kognitiven Fähigkeiten sich mit höherer Wahrscheinlichkeit neidisch verhalten. Das bedeutet, dass sie den Payoff des anderen Kindes entweder immer schmälern oder nur wenn es für sie selbst mit keinen Kosten verbunden ist. Diese Ergebnisse legen nahe, dass ein gewisses Maß an kognitiven Fähigkeiten Voraussetzung ist für neidisches Verhalten, da dieses Verhalten sowohl inhibitorische Kontrolle als auch die Fähigkeit Kosten und Nutzen des Verhaltens abzuwägen, verlangt. Ein weiteres interessantes Ergebnis der Studie ist, dass vor allem unter Jungen mit höheren kognitiven Fähigkeiten neidisches Verhalten sehr ausgeprägt ist. Das bedeutet, dass intelligentere Jungen eher bereit sind, auf eigenen Payoff zu verzichten um das andere Kind in der Allokation schlechter zu stellen als sich selbst. Eine mögliche Erklärung für dieses Verhalten wäre, dass Jungen im Vergleich zu Mädchen kompetitiver sind. Unterschiede in der Wettbewerbsneigung zwischen Mädchen und Jungen wurden schon für Vorschulkinder aufgezeigt und sind möglicherweise der Grund dafür, dass Jungen stärker auf ihre relative Position zu einem anderen Kind achten.

Die zweite Studie, *„Is it the family or the neighborhood? Evidence from sibling and neighbor correlations in youth education and health”* lässt sich inhaltlich in der Literatur zu Chancengleichheit und sozialer Mobilität verorten. Dabei legen wir den Fokus der Analyse auf die Bildung und Gesundheit von Jugendlichen, da diese zwei äußerst wichtige Faktoren bei der Entstehung von Humankapital sind. Die Familie und die Nachbarschaft spielen eine wichtige Rolle für Bildungs- und Gesundheitsoutcomes von Jugendlichen. In dieser Studie untersuchen wir, welcher dieser beiden Einflussfaktoren der bedeutsamere ist um Ungleichheiten zwischen Jugendlichen zu erklären. Um dieser Frage nachzugehen, reichern wir die Survey Daten des SOEP mit administrativen Daten der Bundesagentur für Arbeit auf kleinräumiger Ebene an. Die Verbindung dieser beiden Datenquellen resultiert in einem besonderen Datensatz der es möglich macht, die Wirkmechanismen des Nachbarschaftseinflusses genauer zu untersuchen. Das ist eine besondere Ergänzung bisheriger Studien die entweder personenbezogene Daten oder aggregierte kleinräumige Daten verwendeten, diese aber nicht kombinierten. In einem ersten Schritt berechnen wir Geschwisterkorrelationen in Schulnoten, kognitiven Fähigkeiten, physischer und mentaler Gesundheit sowie BMI von 17-Jährigen. Geschwisterkorrelationen sind ein Maß für den relativen Einfluss familienbedingter Faktoren, da sie angeben wie viel der Varianz in den Outcomes durch Faktoren erklärt werden kann, die von den Geschwistern geteilt werden (z.B. Familie und Nachbarschaft). Ist die Korrelation hoch, so bedeutet das, dass individuelle Faktoren der Jugendlichen keine große Rolle spielen, relativ zum Familien- oder Nachbarschaftshintergrund. Daher liegt in diesem Fall wenig Chancengleichheit unabhängig von der Herkunft vor. Andere Studien haben Geschwisterkorrelationen zwischen 0.4. und 0.6 für Bildungsoutcomes in Norwegen, Schweden und Großbritannien berechnet. Wir schätzen für Deutschland Geschwisterkorrelationen in Bildungsoutcomes zwischen 0.17 und 0.46, also schwächer als in anderen Ländern. Analog zu Geschwisterkorrelationen berechnen wir Korrelationen zwischen benachbarten Jugendlichen, um den relativen Einfluss der Nachbarschaft auf Bildungs- und Gesundheitsoutcomes von Jugendlichen zu schätzen. Empirische Studien haben gezeigt, dass Nachbarkorrelationen verschiedener Arbeitsmarkt- und Bildungsoutcomes geringer sind als Geschwisterkorrelationen. Dies lässt auf den größeren Einfluss familienbedingter Faktoren schließen. Für Norwegen, Schweden und Großbritannien wurde zum Beispiel gezeigt, dass Nachbarkorrelationen für Bildungsoutcomes zwischen 0.02 und 0.14 liegen. Unsere Schätzungen ergeben Nachbarkorrelationen zwischen 0.04 und 0.22. Die mentale und physische Gesundheit betreffend, berechnen wir Nachbarkorrelationen zwischen 0.05 und 0.10. Wir zeigen daher, dass auch in Deutschland Nachbarkorrelationen geringer sind als Geschwisterkorrelationen, d.h. familienbedingte Faktoren wichtiger sind bei der Erklärung von Ungleichheiten zwischen Jugendlichen. Abschließend nutzen wir das gesamte Potenzial unseres Datensatzes, um mögliche Kanäle des Nachbarschaftseinflusses aufzudecken. Dazu wählen wir eine neuartige Herangehensweise der Dekomposition der Nachbarschaftskorrelationen durch regionale Informationen auf kleinräumiger Ebene. Jedoch zeigt sich, dass die uns zur Verfügung stehenden kleinräumigen Informationen auf Postleitzahlebene die Nachbarkorrelationen nicht erklären können.

Zusammenfassend lässt sich sagen, dass unsere Studie empirische Evidenz für den relativen Einfluss von Familie und Nachbarschaft auf das Humankapital von Jugendlichen liefert. Die Frage welcher dieser beiden Einflussfaktoren wichtiger ist, lässt sich klar beantworten – die Familie ist eindeutig bedeutsamer. Darüber hinaus leistet diese Studie einen Beitrag zur bisherigen Literatur, indem sie Nachbarkorrelationen für Gesundheitsoutcomes präsentiert. Diese liefern neue Erkenntnisse, denn sie zeigen, dass die Nachbarschaft vor allem einen Einfluss auf die mentale Gesundheit von Jugendlichen hat. In einem neuartigen Ansatz versuchen wir darüber hinaus den Einfluss der Nachbarschaft zu erklären, können aber mit den Daten die uns zur Verfügung stehen keine Wirkmechanismen offenlegen.

Das dritte Papier trägt den Titel *"Who lies and to whom? Experimental evidence on cheating in a household survey context"*. Darin untersuche ich Präferenzen für ehrliches Verhalten in einer heterogenen Stichprobe von Erwachsenen die Teil des SOEP Innovationssample ist. Ähnlich wie im ersten Papier nutze ich die Stärken beider Datenquellen und setze das Verhalten im Experiment mit sozio-ökonomischen und sozio-demographischen Informationen der Teilnehmer in Beziehung. Darüber hinaus analysiere ich die in dem Datenerhebungsprozess gesammelten Paradaten, um ehrliches Verhalten in dem Experiment näher zu beleuchten. Die Paradaten liefern wichtige Erkenntnisse für die Bedeutung der Umweltbedingungen für ehrliches Verhalten, die bis jetzt noch nicht untersucht wurden. In dem Experiment werden die Teilnehmer aufgefordert einen zweifarbigen Würfel zu werfen und das Ergebnis dem oder der Interviewer/in mitzuteilen. Würfeln sie rot, so bekommen sie entweder einen Euro oder fünf Euro, abhängig davon in welchem Treatment sie sind. Würfeln sie schwarz, gehen sie leer aus. Die Interviewer, die mit den Teilnehmern vor dem Experiment die Standard SOEP Fragebögen durchgehen, wissen nicht, ob die Teilnehmer den Ausgang des Würfelwurfs wahrheitsgemäß berichten oder nicht. Allerdings ist bekannt, dass die Wahrscheinlichkeit eine der beiden Seiten zu würfeln 0.5 beträgt. Vergleicht man 0.5 mit den berichteten aggregierten Wahrscheinlichkeiten, so wird offensichtlich, dass deutlich mehr als 50% - nämlich 57% der Teilnehmer angeben rot gewürfelt zu haben. Vergleicht man die Ergebnisse der beiden Treatments zeigt sich, dass die Versuchspersonen mehr lügen, je höher die Anreize für unehrliches Verhalten sind. So behaupten 62% der Personen die „gute" Farbe gewürfelt zu haben, wenn sie dafür 5 Euro bekommen, während nur 53% der Personen lügen wenn sie dafür einen Euro bekommen. In einem nächsten Schritt vergleiche ich die angegeben Würfelergebnisse in verschiedenen sozio-ökonomischen und sozio-demographischen

Subgruppen. Dabei zeigt sich, dass sich diese kaum unterscheiden zwischen den Gruppen. Eine Ausnahme ist Bildung – je höher die Bildung, desto niedriger ist die Wahrscheinlichkeit zu lügen in dem Experiment. Weit wichtiger für die Erklärung von Verhalten in dem Experiment sind jedoch die Variablen die die Umweltbedingungen beschreiben. Vor allem zwischen den Charakteristika der Interviewer und ehrlichem Verhalten besteht ein hoch signifikanter Zusammenhang. Weibliche Interviewer und ältere Interviewer werden signifikant häufiger angelogen in dieser Stichprobe. Darüber hinaus spielt auch die Dauer des Interviews eine wichtige Rolle; je länger Interviewer und Versuchsperson miteinander agieren, desto niedriger ist die Wahrscheinlichkeit dass die „gute“ Farbe berichtet wird. Diese Ergebnisse deuten auf eine Art der Diskriminierung von Frauen und älteren Personen hin. Außerdem sind die Ergebnisse dieser Studie ein Hinweis darauf, dass Umweltfaktoren eine wichtige Rolle spielen für das Verhalten in Experimenten die Präferenzen für Ehrlichkeit untersuchen. In zukünftigen Studien zu diesem Thema sollten daher Informationen zu der Person mit der interagiert wird nicht außer Acht gelassen werden.

Vorveröffentlichungen

Anmerkung der Autorin

Die folgende Liste enthält alle Vorveröffentlichungen. Darunter sind auch Versionen der Kapitel, die zum Teil stark überarbeitet wurden, bevor sie Eingang in die vorliegende Dissertation fanden.

Kapitel 2: Spite and cognitive skills in preschoolers (mit C. Katharina Spieß)

- Bügelmayer, E., & Spieß, C. K. (2014). Spite and cognitive skills in preschoolers. *Journal of Economic Psychology*, *45*, 154-167.
- Bügelmayer, E., & Spieß, C. K. (2011). Spite and cognitive skills in preschoolers. *SOEPpaper*, (404).

Kapitel 3: Is it the family or the neighborhood? Evidence from sibling and neighbor correlations in youth education and health (mit Daniel D. Schnitzlein)

- Bügelmayer, E., & Schnitzlein, D.D. (2014). Is it the family or the neighborhood? Evidence from sibling and neighbor correlations in youth education and health. *SOEPpaper*, (716).

Kapitel 4: Who lies and to whom? Experimental evidence on cheating in a household survey context (kein Ko-Autor)

- Keine Vorveröffentlichungen

Erklärung gem. § 4 Abs. 2

Hiermit erkläre ich, dass ich mich noch keinem Promotionsverfahren unterzogen oder um Zulassung zu einem solchen beworben habe, und die Dissertation in der gleichen oder einer anderen Fassung bzw. Überarbeitung einer anderen Fakultät, einem Prüfungsausschuss oder einem Fachvertreter an einer anderen Hochschule nicht bereits zur Überprüfung vorgelegen hat.

Ich erkläre außerdem, dass ich meine Dissertation selbstständig verfasst habe.

Berlin, den 16. Dezember 2015
Elisabeth Bügelmayer