The Social Context of Health

Lars Thiel

Wuppertal 2016

The Social Context of Health

Lars Thiel

Inaugural-Dissertation
zur Erlangung des akademischen Grades eines Doktors der
Wirtschaftswissenschaft (Dr. rer. oec.)
an der Fakultät für Wirtschaftswissenschaft
der Bergischen Universität Wuppertal

vorgelegt von
Lars Thiel
aus Perleberg

Wuppertal, den 8. Juni 2016

Die Dissertation kann wie folgt zitiert werden:

urn:nbn:de:hbz:468-20170118-121059-8
[http://nbn-resolving.de/urn/resolver.pl?urn=urn%3Anbn%3Ade%3Ahbz%3A468-20170118-121059-8

Erstgutachter: Prof. Dr. Hendrik Jürges
Zweitgutachterin: Prof. Dr. Kerstin Schneider
Tag der mündlichen Prüfung: 7. Dezember 2016

Danksagungen

Auch die vorliegende Dissertation kommt nicht ohne einen kurzen, aber herzlichen Dank an die nachfolgenden Personen aus, ohne deren Mithilfe die Fertigstellung dieser Arbeit nicht möglich gewesen wäre.

An erster Stelle danke ich meinem Doktorvater, Herrn Prof. Dr. Hendrik Jürges, für seinen persönlichen und fachlichen Beistand. Im Rahmen unserer gemeinsamen Gespräche regte er mich stets zum weiteren Nachdenken und kritischen Hinterfragen an. Ich habe unsere Besprechungen immer als ermutigend und motivierend empfunden.

In diesem Zusammenhang muss auch Frau Prof. Dr. Kerstin Schneider genannt werden, die mich durch die freundliche Übernahme des Zweitgutachtens sowie durch konstruktive Fragen und Anmerkungen während der Disputation unterstützte.

Bei den übrigen Mitgliedern der Promotionskommission, Herrn Prof. Dr. André Betzer, Herrn Prof. Dr. Werner Bönte und Frau Prof. Dr. Juliane Köberlein, bedanke ich mich für die anregende Diskussion während der mündlichen Prüfung.

Einen wertvollen Beitrag zum Gelingen der Dissertation leisteten auch die Teilnehmer auf einer Vielzahl von Forschungskolloquien und Konferenzen, indem sie meine Forschungsergebnisse stets kritisch diskutierten und kommentierten. Hervorheben möchte ich insbesondere Herrn Prof. Dr. Heinz Rothgang und Simon Reif, die durch ihre konstruktiven Koreferate zum Erfolg der Dissertation beigetragen haben.

Meinen aktuellen und ehemaligen Kollegen am Lehrstuhl und dem Bergisches Kompetenzzentrum für Gesundheitsökonomik und Versorgungsforschung danke ich nicht nur für Ihren fachlichen Rat, sondern auch für Ihre moralische Unterstützung und die schöne Zeit außerhalb des Universitätsbetriebs. Vor allem meine langjährigen Kolleginnen Sophie Meyer und Marianna Schaubert waren mir eine wichtige Stütze.

An dieser Stelle sei auch Frau Tanja Bruditz vom Sekretariat insbesondere für ihre Hilfestellung bei organisatorischen und administrativen Fragen herzlich gedankt.

Viel zu verdanken habe ich zuguterletzt natürlich auch meinen Eltern, Edda und Thomas Thiel, meiner Schwester Charlene sowie meinen engsten Freunden, die mich auf meinem Weg durchs Studium und der Promotion begleitet und stets vorbehaltlos unterstützt haben.

Contents

List of Tables

List of Figures

CHAPTER 1

Introduction

1.1 Motivation

This thesis deals with the wider social determinants of health status. Besides genetic factors and personality, socioeconomic status plays a dominant role in predicting health. A large literature has shown that the many facets of socioeconomic status, such as income, education, or occupational level, are strongly related to individual health outcomes (Cutler et al., 2011). In economics, the empirical analysis of individual health determinants and outcomes is predominantly guided by the canonical health demand model developed by Grossman (1972). The standard Grossman model assumes that health production, or health status, is largely influenced by the choices and characteristics of the individual. Hence, he or she chooses the quantity and quality of health-related behavior, health care utilization or drug consumption according to an optimization procedure which leaves out social context effects.

The picture of the social determinants of health, however, is more complex than most empirical studies resting on the health demand model suggest. The social environment in which the individual lives may just be as important as individual actions. There is a large and growing interdisciplinary literature documenting the importance of social context for subjective well-being and health. In their seminal study, Helliwell and Putnam (2004) collect compelling evidence for the conjecture that strong social connections, civic engagement and a trustful environment are powerful and independent predictors of subjective well-being. Epidemiological and economic studies have established a strong link between indicators of individual-level social capital, such as social activities or trust, and health outcomes (for comprehensive reviews, see Kawachi et al., 2008; Folland and Rocco, 2014). The social dependence of health is also mirrored in the large literature dealing with the health effects of relative socioeconomic position. This research is guided by the idea that not only absolute individual income matters for health, but also how one's own income performs relative to some reference point or comparison standard, usually measured as the average income of peers (e.g. Eibner and Evans, 2005; Jones and Wildman, 2008; Gravelle and Sutton, 2009). Thus, a relative disadvantage compared to individuals with similar characteristics can worsen health. This is because having less income than others in a community or society can create psychosocial stress which leads to a higher probability of contracting a disease (e.g. Deaton, 2003; Wilkinson and Pickett, 2006). Finally, research from sociological and psychological studies shows that individuals do not act or evaluate their situation in isolation (for a review, see Thoits, 2011). Hence, a high degree of social integration makes the individual more responsive to the characteristics and actions of others in their social environment. A variety of channels have been proposed through which social relationships affect health status. These include, among others, the dissemination of information on behavioral standards and norms, social comparison processes, or social support (e.g. Berkman et al., 2000; Christakis and Fowler, 2007; Cohen and Wills, 1985).

The social context of individual well-being has also found its way into economic models of individual decision-making and behavior. Social influences have usually been couched in terms of interdependent preferences. Such preferences play a central role in the early works of Veblen (1899) and Duesenberry (1949). They argue that individual consumption decisions are to some extent driven by the consumption level of others. Consumption or income is then a positional good, which produces negative externalities by reducing the utility of those who have less of this good (e.g. Frank, 2008). Interdependent preferences, however, are not only confined to income or consumption. The choices and actions of social ties, or peers in particular, may directly influence individual behavior and well-being in a variety of domains, such as health behavior, education, or crime. Interestingly, the behavior of individuals in the context of social interactions can be modeled as a rational response to the externalities produced by others (e.g. Durlauf and Ioannides, 2010; Manski, 2000). It is thus somewhat surprising that only few attempts have been made to explicitly assess the implications of social interactions for health production and demand, respectively. Notable exceptions are the studies by Bolin et al. (2003) and Folland (2008). Bolin et al. (2003) develop a health demand model where the family is the producer of both health and social capital. Their model predicts strong complementarities between health capital and social capital. On the one hand, a higher stock of both health capital and social capital increases utility. On the other hand, social capital improves the efficiency of the health-production process, for example due to better (shared) knowledge. Folland (2008) includes social capital as an additional input factor in the individual's health-production function. His model predicts a positive correlation between social interactions and the demand for health care goods and services, as long as the health benefits of the larger social capital stock outweigh its consumption benefits.

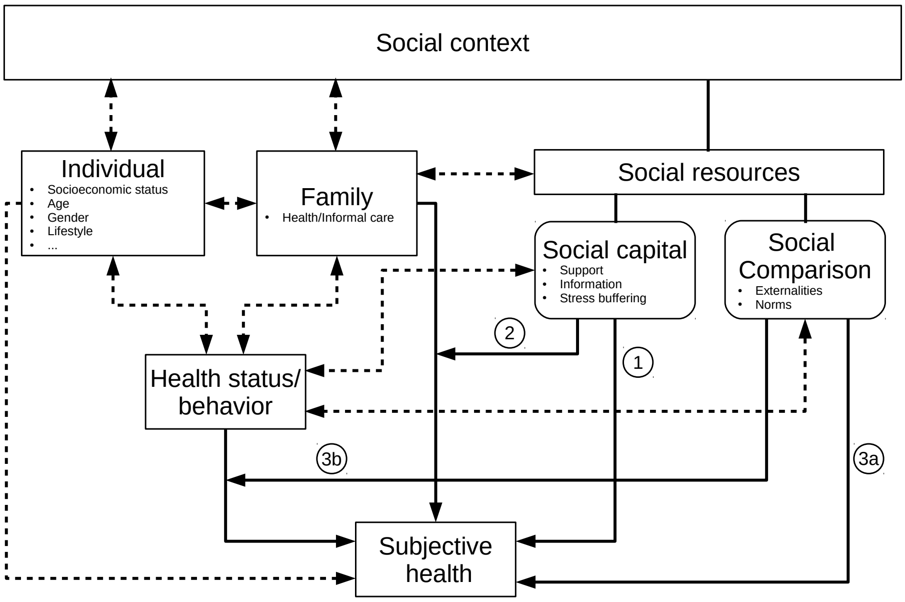
Gaining insights into the importance of social interactions for health is highly relevant from a policy perspective. The accumulating evidence on the social dependence of health suggests that the determinants of health inequalities go beyond socioeconomic factors. Health policy has traditionally been guided by the strong empirical pattern that social inequalities based on income or education are an important predictor of health outcomes (for European evidence, see for example Mackenbach et al., 2008). This has led to efforts to promote the health among the poor and disadvantaged, for instance via redistributing income or facilitating access to health care information and services (e.g. Adler and Newman, 2002). The aim of this thesis is not to downplay the role of socioeconomic gradients in health and the interventions to curb them. It rather proposes that policy makers should put greater emphasis on social context and social interactions as an important determinant of health, which concerns all individuals along the socioeconomic distribution. This thesis is thus in line with the conclusion of some observers that the wider social determinants of health inequalities are rather underrepresented in current health policy debates (e.g. Costa-Font and Mladovsky, 2008; Eckersley, 2015).

Certainly, better knowledge about the impact of social interactions can improve policy makers' understanding about the effectiveness of health policy interventions. As the growing evidence suggests, social ties have an independent effect on a variety of health outcomes. Thus, efforts to strengthen social participation may be a cost-effective strategy to improve population health (e.g. Umberson and Montez, 2010). Research also shows that social networks can facilitate the dissemination of health-related norms and outcomes, for instance in the context of body shape or mental well-being (Christakis and Fowler, 2007; Fowler and Christakis, 2008). This indicates that health-promoting measures have spillover or multiplier effects that should be taken into account in cost-effectiveness analyses. Neglecting such externalities may bias cost-effectiveness estimates and lead to a non-optimal allocation of resources within the health care sector (e.g. Labelle and Hurley, 1992). What is more, the presence of health-related externalities may influence the decision whether to increase public spending on health care. For example, individuals may have a preference for public provision of health care goods and services when they have altruistic preferences and care about other people's health status (e.g. Paolucci, 2011).

These are just some of the many examples on how the social dependence of health may influence policy decisions. Generally, there are strong theoretical and empirical arguments for the role of social interactions in the health domain. This thesis therefore analyzes the empirical content of specific mechanisms that explain the positive association between social context and health.

1.2 Aim and Contribution

This thesis contributes to the emerging literature on social interactions and individual health outcomes. The focus is not on theoretical modeling. The present thesis fills gaps by explicitly examining previously underinvestigated social determinants of individual health. The aim is to substantiate the conjecture that individual health production hinges on social interdependencies. The conceptual framework of this thesis is illustrated in Figure (1.1). Individuals and their families are assumed to be embedded in social networks which influence their decisions and outcomes. The social environment not only affects health status or behavior, but also how individuals evaluate their health situation (subjective health). The dashed arrows mirror the rather complex interactions between individual characteristics, social context and health. Clearly, individual actions and characteristics and the social environment may mutually affect each other in the health production process. The studies in this thesis analyze the influence of social context on health taking these interrelationships as given. An important presumption is that



Source: Own illustration.

Figure 1.1: Conceptual framework of the thesis

social context influences health status and subjective health through social resources that can be built up by social participation.

The first type of resources analyzed in this thesis can be subsumed under the umbrella term social capital. There are a variety of mechanisms that can lead to a positive relationship between social capital and health. Frequent social interactions may serve as a source of distraction and social support. For instance, social contacts may help with emotional problems or provide practical assistance (e.g. Thoits, 2011). Stronger social ties may also improve an individual's knowledge about the workings of the health care system and facilitate access to medical services and treatments (e.g. Folland, 2008). Furthermore, the resources embedded in social capital may facilitate coping with demanding situations. Thus, stronger social integration may help individuals deal with stressful situations, which has been coined as the stress-buffering property of social capital (e.g. Kawachi and Berkman, 2001). As indicated by the solid arrows 1 and 2, social capital thus has both a direct and a moderating influence on health. The direct association between social capital and health is the subject of Chapter (2). Specifically, it examines the relationship between cultural participation and self-rated health. Cultural activities are an important aspect of an individual's social life and social capital, but their health implications have rather been underexplored. The moderating role of social capital is analyzed in Chapter (3). It deals with the effect of caring for a loved one at home on mental health, and how the adverse psychological consequences of caring are influenced by the individual's level of social capital and participation, respectively.

The second type of social resources analyzed in this thesis arises from social comparisons. Frequent social interactions make the individual susceptible to the actions and characteristics of others. In the absence of an objective comparison standard, many individuals use the performance or attributes of similar others to evaluate their situation (Festinger, 1954). This is particularly relevant in the health domain, where individuals frequently turn to external reference points provided by stereotypical others (e.g. Groot, 2000). From an economic perspective, this creates externalities which influences individual well-being and health assessment (solid arrow 3a). On the one hand, the health-related spillovers can be positive if individuals care about other people's health status (e.g. Culyer and Simpson, 1980). On the other hand, the social comparison effect on health is negative when similar others are healthier and may thus have a relative advantage (e.g. Mujcic and Frijters, 2015). The health of others may also influence subjective health assessment through altering the perception of own health status (solid arrow 3b). The health of peers thus acts as a perceived social norm that need to be followed. Falling short of this standard may lead to psychological costs and actually aggravate the negative effect of poor health status on subjective health (e.g. Powdthavee, 2009). Chapter (4) examines the associations between health-related social comparisons and individual health evaluation. We introduce a relative health variable which measures the average health status within a respondent's social reference group. It is assumed that individual health satisfaction not only depends on own health status, but also on how own health performs relative to the medical condition of peers.

Generally, the three studies make the following common contributions. First, the present thesis provides novel evidence regarding the relationship between social context and health for Germany. The vast majority of empirical studies dealing with the type of social interactions analyzed in this doctoral thesis have used data from the United States, the United Kingdom or Scandinavian countries. However, insights from these countries can neither be compared, nor applied to different institutional and cultural contexts. Although the evidence is not conclusive, institutional differences are important in shaping individual-level social networks. It is, for instance, possible that a larger welfare state may suppress social interactions (e.g. Van Oorschot and Arts, 2005). As a consequence, individuals might be less responsive to the benefits (and harms) of social relationships. Furthermore, it is also possible that a more generous social security system crowds out informal support obtained from social contacts (e.g. Clark and Lelkes, 2006). Thus, to further corroborate the evidence, it is useful to analyze the importance of social influences on health in a different institutional and cultural context.

Second, this thesis relies on large-scale population level data. Previous studies on the topics discussed here have used rather small and non-representative samples. The empirical analyses in the present study rely on survey data from the German Socio-Economic Panel (SOEP) (for an overview, see Wagner et al., 2007). It is the richest source of individual-level data over a longer period of time in Germany. The core of the survey consists of measuring the socioeconomic and demographic conditions of the German population. However, the questionnaire has been extended over the years to include indicators for health status, health behavior, health-related quality of life, leisure and social activities, personality traits, and retrospective information on conditions and socialization in childhood and youth.

Third, the empirical estimates presented in this thesis go beyond simple associations. The estimates of previous studies are potentially biased, because many characteristics that influence social interactions and health have been omitted. In contrast, the empirical analyses included in this thesis take a rich set of control variables into account, which should reduce problems arising from omitted variables bias and unobserved heterogeneity, respectively. Bias is further reduced by exploiting the longitudinal nature of the data. In the SOEP survey, individuals are followed over an extended period of time. This feature can be used to eliminate the effects of unobserved factors that do not change over time, such as genetic make-up, preferences, personality, or motivation. This has been done in Chapter (3) and (4), which estimate fixed effects models and rely on changes of dependent and independent variables over time. In Chapter (2), the empirical analysis exploits longitudinal, or rather lagged, information on cultural participation and health to reduce concerns that the findings are driven by health-related selection and confounding due to unobserved preferences and skills.

Clearly, the empirical approach adopted in this thesis does not completely solve the problem of unobserved heterogeneity. Although residual bias due to omitted variables could remain, this thesis provides a major improvement compared to previous studies which have not explicitly addressed the problems due to unobserved selection. However, each contribution discusses the conditions under which the obtained estimates are biased and cannot be interpreted as causal.

1.3 Overview

Chapter (2): Cultural Participation and Health

The study in Chapter (2) deals with the direct implications of social activities for individual health. Instead of looking at social participation in general, this investigation focuses on a specific type of social activity – which is cultural participation or attendance. The consumption of cultural activities is an integral part of an individual's social life. A survey on further training and leisure activities in the adult German population shows that about 60 percent report that they visited cultural events or consumed cultural activities (theater, opera or classical concerts) in the last 12 months (Authoring Group Educational Reporting, 2012). The relevance of the cultural sector in Germany is also reflected in high public expenses. In 2009, approximately 9 billion euros (about 0.4 percent of GDP) were spent for culture and affiliated areas. The public provision of cultural goods and services can be justified by the positive effects of arts participation on human capital development and subjective well-being. Theoretically, frequent exposure to arts activities facilitates the acquisition of arts-related human capital and skills (Stigler and Becker, 1977). Moreover, a number of empirical studies has shown that cultural activities are positively related to individual outcomes such as educational attainment, cognitive skills, social capital and quality of life (e.g. DiMaggio, 1982; Hille and Schupp, 2015; Jeannotte, 2003; Kim and Kim, 2009).

Positive health effects of cultural attendance may arise due to various reasons. Frequent consumption of artistic activities may create a mentally stimulating environment that facilitates cognitive competencies (Hertzog et al., 2008; Stine-Morrow et al., 2007). Cultural participation may also serve as a coping mechanism in the context of stressful situations or negative life events (e.g. Iwasaki et al., 2005). What is more, participating in cultural activities may simply be an expression of high socioeconomic status. This may reinforce social hierarchies, leading to positive health effects for the better-off (Bourdieu, 1984; Wilkinson, 1999). The positive health effects of cultural participation can also be ascribed to the benefits of social capital or social interactions. According to Folland (2008), social capital improves health because social networks provide resources for stress buffering, serve as a source of information, and may increase the sense of responsibility towards oneself and others in the community.

Many studies have found a positive correlation between cultural participation and health outcomes (e.g. Cuypers et al., 2012; Johansson et al., 2001; Khawaja and Mowafi, 2006; Renton et al., 2012; Wilkinson et al., 2007). However, a key limitation of this research is that it does not address the problem of endogeneity. The positive and large health benefits could simply reflect selection of healthy individuals into cultural attendance. Furthermore, unobserved variables that simultaneously influence arts participation and health could bias the estimates. Major candidates for unobserved confounders are preferences and skills related to the arts (Ganzeboom, 1984). These traits are usually unobserved but can strongly influence the selection process. Furthermore, these traits may be acquired at birth or due to early and frequent exposure to the arts over the life cycle. This creates a further enodgeneity problem, because contemporaneous cultural participation, health and other background characteristics correlate with past cultural activities via unobserved preferences or skills (Throsby, 1994).

The major aim of this study is to provide a thorough empirical analysis using the SOEP data to assess the association between cultural participation and health. The cornerstone of the empirical approach is a propensity-score matching procedure, to obtain a homogeneous sample of cultural participants and non-participants. To reduce potential imbalances between these two groups, individuals are matched conditional on a rich set of control variables. It includes many variables that have largely been omitted in previous studies, such as social and leisure activities, health-related lifestyle, personality traits, and early exposure to the arts. To further alleviate concerns that significant differences between individuals remain, we perform a regression analysis based on the matched sample, with health as the dependent variable and cultural participation and all covariates as independent variables. We tackle the problem of reverse causality by including a lagged health outcome among the control variables. Furthermore, concerns about unobserved heterogeneity are reduced by performing the matching analysis in subsamples of individuals defined by the same level of cultural participation one year before the effect of cultural participation is measured. This approach reduces residual bias due to unobserved selection into cultural attendance and mitigates the problem of endogenous control variables (Lechner, 2009).

After taking simultaneity and omitted variables bias into account, the empirical analysis reveals a positive and significant effect of regular cultural participation on mental health only. The mental health benefit is generally robust to variations of the matching procedure. To assess under which conditions the positive psychological gain becomes zero, we perform a sensitivity analysis that has been developed by Ichino et al. (2008) specifically for matching estimators. Basically, we investigate the responsiveness of the matching estimate to the inclusion of a simulated confounder variable U in the conditioning set. This variable can be defined as having the same empirical distribution as observed characteristics. Another possibility is to search for an empirical distribution that renders the matching estimate zero. We perform both exercises.

We find that the greatest threat to the positive mental health effects of cultural participation comes from unobserved factors that strongly influence selection into cultural attendance, and improve mental health even in the absence of arts activities. What is more, by performing a sensitivity analysis similar to Altonji et al. (2005), we find that an unobserved confounder, which has similar selection and outcome effects as some observed covariates, almost halves the estimated mental health benefit.

The empirical results raise some doubt about the effectiveness of cultural policy to reduce health inequalities. Despite widespread public funding of arts and cultural institutions, arts participation is still more prevalent among higher social status or education groups (Authoring Group Educational Reporting, 2012). One explanation could be that the supply of arts performances predominantly reaches those from rather advantaged socioeconomic backgrounds. These individuals may have acquired arts-related preferences and skills, for instance due to early exposure to the arts and parental infuences, or frequent consumption of arts activities.

Chapter (3): Social Capital, Caregiving and Mental Health

While Chapter (2) deals with the direct influences of social activities on health, the study in Chapter (3) looks at the moderating role of social interactions, or social capital, among individuals under demanding situations. The *stress-buffering hypothesis* says that social networks can alleviate the negative health implications of stressful events and tasks (Cohen and Wills, 1985; Kawachi and Berkman, 2001).

Specifically, the analysis focuses on informal caregiving as a presumably demanding task. Caring for a loved one can be an extraordinarily stressful situation, because it is usually a very time-consuming activity which interferes with other duties such as work, leisure and other family activities. The economic costs of informal caring can be severe. Caregivers might themselves get ill, which increases sickness absence rates and raises the demand for health care and formal long-term care services (e.g. Bauer and Sousa-Poza, 2015). Informal care is often associated with a high psychological burden. Empirical research suggests that informal caregivers frequently report depressive symptoms, anxiety, or stress and exhibit low levels of subjective well-being (e.g. Schulz and Sherwood, 2008). Developed countries have acknowledged the problem of caregiver strain and have introduced and expanded policies that are specifically directed at the well-being of informal carers. For instance, the German health insurance and long-term care providers offer respite care, training and counseling, and coordinated information services for family carers. However, many insured persons do not use these services (Robert Koch Institute, 2015).

Alternative sources of support may come from the individual's social network or social capital (see also Thoits, 2011). Social ties can reduce the psychological burden of informal care by providing emotional support. Moreover, caregivers may receive practical or financial assistance from social connections. Social contacts are also an important resource of information for the caregiver. Significant others may, for instance, provide knowledge about the availability and accessibility of public caregiver support services. In accordance with the stress-buffering hypothesis, a higher degree of social capital should thus moderate the negative association between informal care and mental health.

In the empirical analysis, the buffering effect is tested by including an interaction term between a caregiving indicator variable and a social capital index, which basically represents the weighted sum of social activities. The results show that providing non-zero hours of care to a household member negatively influences mental health. However, the negative association is attenuated by a higher level of social capital and social participation, respectively. Furthermore, individuals who report high caring intensities suffer more, but also experience the strongest buffering effects by social capital. Finally, those who volunteer regularly and provide informal care at the same time experience the largest buffering effects compared to other social activities.

A number of sensitivity checks suggests that the buffering effect of social capital is robust to different explanations of the buffering effect. The buffering effect of actual social capital persists even if we control for the moderating role of observed covariates correlated with social activities. Furthermore, caregiving context, such as the health of the care recipient or other sources of help, cannot explain why carers with more social capital are healthier than carers with less social capital. However, when we use a matched sample to reduce the bias in observable characteristics between carers and non-carers, the coefficient for the moderator effect decreases substantially.

Nevertheless, problems due to the endogeneity between caring and social activities remain. It is possible that the buffering effect is overestimated and that it is actually driven by the potential simultaneity between caring decisions and social activities.

Chapter (4): Health Satisfaction and Relative Health

The study in Section (4) analyzes the role of relative concerns in the context of health status for individual health satisfaction. Frequent social interactions may increase the susceptibility of individuals to the characteristics of others. In many cases, it is important how one performs relative to others, and social contacts frequently act as a reference point for comparison. This reasoning is prominent in the literature dealing with the effect of income on happiness and health status. Subjective well-being is thus not only influenced by absolute individual income, but also by how it performs relative to the income of peers (e.g. Clark et al., 2008; Wagstaff and Van Doorslaer, 2000).

Such relative comparisons might also be relevant in the health domain. The health status of others potentially influences how own health is evaluated. A leading example from the health economics literature is the scale of reference bias in subjective health and quality of life assessments. When asked to judge their own health status, respondents tend to compare their health with the condition of similar other people (e.g. Fienberg et al., 1985). This poses problems for the use of self-assessed health measures in surveys or cost-effectiveness analyses. Due to different standards of comparison, subjective health assessments may not be interpersonally comparable. Individuals with objectively the same health status may have different health perceptions, whereas respondents who perceive their health status equally may have different levels of true health (e.g. Groot, 2000).

The existence of health-related externalities provide further support for relative health effects on individual well-being and health evaluation, respectively. The canonical example of external health benefits or positive health spillovers is the reduced risk of infection when others are vaccinated against communicable diseases (Culyer, 1971). However, the health status of others per se might influence subjective well-being. This is because individuals may not only care about other's well-being or utility, but also about other persons' health status (Culyer and Simpson, 1980). These caring externalities can be explained by altruistic preferences. Individuals are thus willing to sacrifice own resources and health to improve other people's health status (e.g. Hurley and Mentzakis, 2013).

However, the health status of peers can also produce negative externalities. This is because the relative standing in the societal health distribution may involve social status effects. Good health provides both economic and non-economic effects. Besides psychological and social benefits, a relative advantage in the health domain can provide tangible benefits on marriage and labor markets (Mujcic and Frijters, 2015). For instance, those with better relative health may find superior marriage partners in terms of health status and socioeconomic position (e.g. Wilson, 2002). Relative health status may also determine labor market success. There is evidence that good health, and also physical fitness and attractiveness, is positively related with the chance of having a prestigious job and high earnings (e.g. Hamermesh and Biddle, 1994; Lindeboom, 2006).

This study empirically tests whether and how relative health affects individual health evaluation and satisfaction, respectively. An important contribution is made here, since we include a measure of relative health as a predictor of health satisfaction. Relative health is measured as the average health status in the respondent's reference group, calculated based on an overall health index. The reference group consist of individuals that have the same age, gender, educational level, and live in the same region in a given year. Fixed effects models are used to estimate the association between reference-group health and satisfaction with health, conditional on own health status and covariates. The empirical analysis also explores the possibility that the effect of other people's health may differ depending on whether one is healthier or sicker than the reference point and the peer group, respectively. Research from social psychology suggests that social comparisons, whether they are with sicker or healthier persons, can have both positive and negative psychological effects (Buunk et al., 1990). Finally, this study also investigates how reference-group health changes the association between own health status and health satisfaction. The basic hypothesis is that the health of significant others acts as a social norm or standard, and that a deviation from this standard may bring about psychological costs (Powdthavee, 2009). It is therefore expected that an improvement of reference-group health facilitates the satisfaction effects of good health, or increases the mental burden of poor health status.

The results show that relative health is not strongly associated with individual health satisfaction. Changes in reference-group health are not significantly related to changes in health satisfaction. The empirical estimates suggest that there could be a health-norm effect. However, the point estimates are quantitatively and economically negligible.

The empirical results are actually good news for researchers who use subjective health or quality of life measures, such as satisfaction with health, as outcomes in empirical analyses or cost-effectiveness studies. One could conclude that the bias in subjective health measures due to reference-group effects is negligible and that these measures are interpersonally comparable.

CHAPTER 2

Cultural Participation and Health

2.1 Introduction

There is a growing political and academic debate about the value and impact of culture. Policy makers become increasingly aware of the benefits of the arts, particularly in the context of education and further training activities. In a report on the state of the cultural sector in Germany commissioned by the German parliament, cultural activities have been recognized as a key ingredient to improve individual outcomes, such as education, cognitive skills and health (German Federal Parliament, 2007). The relevance of the cultural sector in Germany, as in many other countries, is also reflected in high public expenses for culture and affiliated areas. In 2009, the subsidies roughly amounted to 9 billion euros (about 0.4 percent of GDP), which were largely spent on theaters and musical arrangements (35 percent) and museums, exhibitions, and collections (18 percent) (Statistical Offices of the Federation and the Länder, 2012). Public funding of arts and culture can on the one hand be justified by economic reasoning. In their seminal paper, Stigler and Becker (1977) argue that frequent exposure to arts activities improves the individual's skills of understanding and appreciating the arts good and facilitates the accumulation of arts-specific human capital, respectively. A different approach by Lévy-Garboua and Montmarquette (1996) assumes that tastes are unknown and revealed to the consumer only when experiencing artistic performances. Exposure to artistic experiences can thus create a negative or positive shock in tastes or preferences for specific types of arts or arts in general. On the other hand, there is ample empirical evidence that cultural activities are positively related to a variety of individual outcomes such as educational attainment, cognitive skills, social capital and quality of life (e.g. DiMaggio, 1982; Hille and Schupp, 2015; Jeannotte, 2003; Kim and Kim, 2009). Hence, increasing the supply of cultural goods and services, and encouraging arts consumption, via public funding may pay off in terms of higher individual well-being and utility, respectively.

The subsidization of the cultural sector may also have implications for public health. A large number of medical studies stresses the role of arts activities in clinical settings (McCarthy et al., 2004). Recent population-level evidence shows that rather passive or receptive arts activities, such as attending artistic performances, are beneficial for perceived and behavioral health outcomes (e.g. Wilkinson et al., 2007; Cuypers et al., 2012; Renton et al., 2012). A key limitation of this research is that it does not seriously address the problem of simultaneity and unobserved heterogeneity. Simultaneity bias arises due to health-related selection or, that is to say, reverse causality. Previous estimates could to a large extent reflect selection of healthy individuals into cultural attendance rather than a causal impact of arts participation on health. Unobserved heterogeneity results from omitted variables. Many factors that influence both cultural participation and health — for instance social and leisure activities, health-related lifestyle, personality, and early exposure to the arts — are either excluded or only imperfectly measured. A spurious positive correlation between arts activities and health may also stem from unobserved elements of preferences and cognitive skills related to the arts, which strongly influence selection into cultural attendance (e.g. Ganzeboom, 1984; Notten et al., 2015). Unobservable taste and skills can also create problems due to endogenous control variables. These traits may be acquired at birth or due to early and frequent exposure to the arts over the life cycle. Thus, it is likely that past cultural activities influence the control variables via unobserved preferences and abilities for the arts (e.g. Throsby, 1994).

The analysis presented in this chapter complements past research by conducting a careful econometric analysis of the association between cultural participation and health. We explore the possibility that previous investigations overestimate the health benefits of arts attendance. Despite including a rich set of covariates in the empirical analysis, we cannot find a robust effect of cultural attendance on self-reported health outcomes. The greatest threat to valid causal inference is the existence of unobservable traits that strongly predict cultural participation and simultaneously improve health outcomes. Controlling for this specific deviation from unconfoundednes may substantially reduce the positive health effects of arts activities.

The empirical analysis relies on rich individual-level data from the German Socio-Economic Panel (SOEP). Cultural participation is measured as regular (at least monthly) consumption of cultural events like opera, theater or lectures. As outcomes, we use summary scales for physical and mental health. The set of conditioning variables contains demographic and socioeconomic characteristics, social and leisure activities, personality traits, markers for health behavior, and indicators for early exposure to the arts (for instance parental education and childhood artistic activity). The econometric analysis starts with a simple linear regression of health outcomes on the cultural participation indicator. We determine how the estimated relationship changes as we add groups of control variables. Generally, when the number of conditioning variables increases, the health benefit of cultural attendance decreases. After conditioning on the full set of covariates, a significant association exclusively with mental health remains.

The remainder of the empirical analysis relies on matching estimates. To this end, we apply a propensity-score matching procedure and obtain a homogeneous sample of cultural participants and non-participants. Linear regressions are then performed on this matched sample to reduce observable differences between these two groups. The matching results including all control variables produces estimation results similar to the linear regressions. To see how the estimates change when we take reverse causality into account, we also match on lagged health outcomes. The findings indicate that healthy individuals select into cultural attendance, which drives the effect of arts participation upwards. We reduce concerns about unobserved heterogeneity by performing the matching analysis in subsamples of individuals with the same level of past cultural participation. This approach reduces the residual bias due to unobserved selection into cultural attendance and mitigates the problem of endogenous control variables (Lechner, 2009). The estimate for mental health decreases but remains statistically significant.

The sensitivity of the matching estimates is assessed by performing a variety of robustness checks. The matching analysis is replicated using various matching algorithms, different conditioning sets, an alternative region of common support, and bootstrapped standard errors. However, applying different methods does not change the baseline estimates substantially. Moreover, and more important, we conduct a sensitivity analysis proposed by Ichino et al. (2008), to assess the robustness of our findings with respect to the inclusion of an unobserved binary variable. Given certain assumptions about the direction of unobserved confounding, we can examine under which conditions the estimated relationship between cultural attendance and mental health is biased upwards. This approach can also be used to provide information about the extent of omitted variables bias needed to reduce the size of the coefficient to zero. This exercise indicates that the positive association between cultural participation and mental well-being may reflect positive selection into cultural engagement due to unobserved factors. Specifically, we find that an unobserved confounder simulated to have outcome and selection effects similar to several observable covariates is enough to reduce the point estimate by roughly 50 percent. Unobserved preferences for the arts or cognitive skills related to arts appreciation heavily influence participation in cultural activities and can overestimate the effect of cultural engagement on health.

The remainder of this chapter is organized as follows: Section (2.2) describes the main mechanisms underlying the arts-health relationship and discusses the empirical evidence on the association of cultural event attendance with health. Section (2.3) outlines the empirical approach adopted in this study. Section (2.4) details the data and estimation sample. Section (2.5) presents and discusses the estimation results. Section (2.6) provides the robustness and sensitivity analyses. Section (2.7) concludes.

2.2 Related Literature

2.2.1 Theoretical Considerations

From a theoretical perspective, there are a variety of mechanisms through which cultural participation can influence health. Frequent visits to cultural events possibly provide a stimulating environment that could lower the rates of cognitive aging and enhance levels of cognitive functioning in old age (Hertzog et al., 2008; Stine-Morrow et al., 2007). Furthermore, cultural activities such as visiting a museum or an opera may be used as a coping strategy to deal with health problems (e.g. Iwasaki et al., 2005). Cultural events may thus provide an opportunity to deal with everyday problems or negative life events improving physical and psychological well-being. Following Abel (2008), cultural activities such as arts attendance may also reflect socioeconomic status, and it is well-known that individuals with higher incomes or better education tend to be healthier than others (for a review, see Cutler et al., 2011). Similarly, Khawaja and Mowafi (2006) argue that cultural activities could reflect social stratification in society. Accordingly, to maintain and accumulate their social status, individuals invest in cultural capital, for example via visits to arts events. Following Bourdieu (1984), this behavior sustains and creates social hierarchies, which could have deleterious health effects at the individual level. This is in line with the hypothesis put forward by Wilkinson (1999), which states that social hierarchies are associated with psychosocial stress, aggressiveness, less trustfulness and lower levels of social cohesion.

The majority of arts activities involves social interactions with other persons that could form the basis of an individual's stock of social capital. Individual-level social capital, such as the frequency and intensity of personal contacts, have been shown to positively influence health and survival (for a review, see Kawachi et al., 2008). Hence, visits to cultural events can positively influence health via the benefits of social networks and interactions (see also Hyyppä, 2010). These include, for instance, stress reduction and the provision of information on how to effectively deal with diseases. Cultural activities can thus be seen as a form of social capital that can be used as inputs in health production (e.g. Folland, 2008). From this perspective, individuals may invest in their cultural capital by increasing their consumption of artistic activities. In return, they could tap on the resources involved in cultural capital, such as social connections, that facilitate coping with health problems.

2.2.2 Empirical Evidence

Several empirical studies using individual-level survey data have found a positive association of cultural attendance with perceived health in various populations (Cuypers et al., 2012; Johansson et al., 2001; Khawaja and Mowafi, 2006; Renton et al., 2012; Wilkinson et al., 2007). It is questionable, however, whether the observed correlations reflect causal effects of arts participation on health. Cultural participation is potentially endogenous due to unobserved heterogeneity and reverse causality. The previous observational studies include a variety of personal characteristics in their regression models to mitigate omitted variables bias, but still leave out or incompletely measure many factors that correlate with both the level of arts participation and health outcomes. These especially include social activities, lifestyle, personality straits, early determinants of arts participation, and preferences and cognitive skills. Moreover, only few studies consider the possibility that health might influence cultural participation.

We will improve upon the previous research by performing regression analyses using a homogeneous (matched) sample of cultural participants and non-participants. The problem of omitted variables bias is mitigated by including a rich set of control variables that accounts for the respondent's social and leisure life, health-related lifestyle, personality and childhood exposure to the arts. We will deal with reverse causality by including lagged health outcomes as covariates. Concerns about residual bias due to unobserved confounders are reduced by conditioning on the levels of past cultural activity. According to Lechner and Sari (2015), since baseline cultural attendance is the same, we mitigate the problem of unobserved third variables affecting arts participation and health outcomes. However, we will also show that this may not completely solve the problem of endogeneity when we allow for certain deviations from the unconfoundedness assumption.

Complementary evidence on the health benefits of arts participation comes from controlled experiments in a small region of Sweden (Bygren et al., 2009; Konlaan et al., 2000). They have estimated the effect of, among others, cultural attendance on medical and self-rated health outcomes using a randomized controlled trial (RCT). The treatment was randomly assigned to the study participants, who were encouraged to increase their cultural participation by being offered a free ticket per week for (highbrow) cultural events. Thus, the estimates presumably suffer less from non-random selection into cultural activities based on unobserved factors and health status. Results suggest that treated individuals perform better with respect to a variety of clinical outcomes and selected aspects related to mental health. However, the findings of these studies are based on highly selective samples and are not applicable to the general population.

2.3 Empirical Approach

The empirical analysis basically proceeds in two stages. In the first stage, we estimate simple linear regression models to assess the relationship between cultural attendance and health. By starting with an unconditional correlation and successively adding groups of covariates, one can assess to what extent the health benefits of arts participation are driven by personal and family background characteristics, some of which have been omitted or unmeasured in previous studies. In the second stage, we use matching for preprocessing the data to obtain a homogeneous sample of cultural participants and non-participants. Matching cannot improve causal inference but it can be exploited to reduce bias in observed characteristics before running the regression analysis (Ho et al., 2007).

The matching procedure adopted in this chapter involves the following steps: First, participants are matched to non-participants based on propensity scores. The individual propensity score corresponds to the predicted probability of (regularly) visiting cultural events conditional on observed covariates, which is estimated using a probit regression. The predicted values from this model are used to reweight observations to create a comparison group that is similar to the group of culturally active individuals in terms of observable characteristics. Various matching algorithms exist to calculate the weights, and the choice often involves a trade-off between bias and efficiency (Caliendo and Kopeinig, 2008). We primarily rely on the kernel matching estimator as proposed by Heckman et al. (1998) and Smith and Todd (2005), and use the Epanechnikov kernel with a bandwidth equal to 0.06. Hence, we compare the outcome of cultural participants with the weighted average outcome of all control units, and more weight is given to individuals with similar propensity scores.[1] Second, we perform a regression on the matched sample, that is we include the weights resulting from the matching procedure in a linear regression model of the outcome on the cultural participation indicator and all control variables.

The method employed in this study provides a number of advantages over simple linear regression models. First, instead of matching on combinations of covariates, the similarity of the respondents is assessed on the basis of one single number, which is the propensity score. According to Rosenbaum and Rubin (1983), in situations where the set of covariates is large and high-dimensional, it suffices to condition on the propensity score. This approach avoids the problem of empty cells when matching treated and control individuals directly based on a large number of covariates. Second, by performing regression adjustment after matching, one is confident to overcome the problems associated with misspecifications of either the propensity-score model or the linear regression, which is often referred to as the double-robustness property of regression-adjusted matching models (e.g. Bang and Robins, 2005). Third, since the regression model additionally conditions on all observed baseline characteristics, we are also able to reduce the bias emanating from remaining covariate imbalances after matching (e.g. Imbens and Wooldridge, 2009). Fourth, regression on the matched sample provides non-parametric estimates since it assumes no functional form between the outcomes and the explanatory variable (Imbens, 2004). This allows the association between cultural participation and health to be non-linear and heterogeneous across individuals, respectively. Fifth, matching enables the researcher to assess the similarity between culturally active and inactive individuals. Careful examination of both the covariate distributions among participants and non-participants and the region of common support ensures that cases and controls theoretically have the same probability of being treated (e.g. Dehejia and Wahba, 1999). The latter is particularly important when there is insufficient overlap between case and control individuals, that is when there are not enough comparable control individuals in terms of the propensity score. In contrast to simple linear regression, matching allows explicit assessment of overlap and avoids such extrapolation, for example by excluding observations with propensity scores equal or close to zero (see Gelman and Hill, 2006, chap. 10).

The mean difference between cultural participants and non-participants is calculated as the *average treatment effect on the treated* (e.g. Smith and Todd, 2005):

[1] The matching weights are estimated using the *psmatch2* program in Stata, which has been developed by Leuven and Sianesi (2003).

$$\widehat{ATT} = \frac{1}{N_1} \sum_{i \in I_1} \left[Y_{1i} - \sum_{j \in I_0} W(i,j) Y_{0i} \right], \tag{2.1}$$

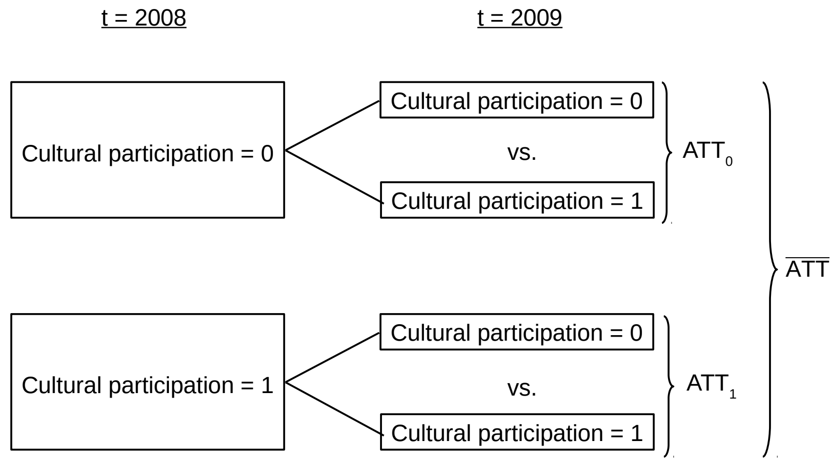
where N_1 is the number individuals in the treatment group (cultural participants) used for the calculation of the ATT, I_1 is the set of treated individuals, I_0 is the set of control individuals (non-participants). $W(i,j)$ is a weighting function that assigns to each control unit a weight calculated based on the distance of propensity scores between the treated and matched control individuals. We use the Epanechniko kernel with a bandwidth equal to 0.06 to construct the weights for matched observations. Thus, cultural participants receive a weight equal to one, and the weight of non-participants is a decreasing function of the distance of propensity scores between treated and control subjects (see also, for example, Heckman et al., 1998).

A critical requirement for unbiased estimates of cultural participation effects is unconfoundedness. A significant and positive association between arts participation and health might be caused by omitted variables that are simultaneously related to health and the decision to attend cultural events. In other words, uncontrolled differences between culturally active and inactive individuals may overstate the benefits of arts attendance, because those who consume artistic activities may have better health in the absence of participation due to unobserved characteristics. We seek to reduce this bias by including a wide array of personal and family background variables (see Section (2.4)).

The observed health benefits of cultural participation could still reflect simultaneity between arts attendance and health. The positive association between cultural attendance and health could just reflect the greater ability of healthy individuals to participate in social and cultural activities. To assess the sensitivity of our estimates to this type of reverse causality, we also control for the health outcome measured before the effect of cultural participation is analyzed.

Nevertheless, it is likely that we incompletely control for unobserved traits related to the arts despite including a rich set of control variables. From an economic perspective, preferences for the arts are endogenously determined and depend on past arts consumption. Frequent and early exposure to the arts increases the utility of future cultural consumption, by loosening budget constraints or altering preferences for the arts (Stigler and Becker, 1977; Lévy-Garboua and Montmarquette, 1996). Variations in individual preferences for the arts have frequently been mentioned as a cause for socioeconomic gradients in arts participation. Individuals with higher social status are more likely to participate in arts activities because, for instance, they were more frequently exposed to the arts by their parents or in schools (e.g. Gray, 1998; Borgonovi, 2004). Differences in cognitive skills may also produce social gradients in cultural participation, particularly with respect to education. Individuals may have different capacities to process cultural information depending on a variety of factors including artistic talent, schooling and knowledge and experience associated with the arts (e.g. Ganzeboom, 1984; Notten et al., 2015) Furthermore, cognitively more able individuals are assumed to be healthier since they are able to process health-related information more efficiently than others (e.g. Auld and Sidhu, 2005).

To reduce concerns about unobserved heterogeneity, we follow the approach proposed by Lechner and Sari (2015). This strategy reduces selection bias by performing the matching analysis within subsamples defined by past cultural participation in the baseline period (here: 2008). This procedure is illustrated in Figure (2.1). Thus in each strata defined by cultural participation in 2008, we basically estimate the mean health difference between active and inactive individuals in 2009 and 2010, respectively. We obtain an aggregate estimate by computing the average of the matching estimates over the two groups, where the coefficients and variances are weighted by the stratum-specific number of observations (see also Schmitz and Westphal, 2015). By conditioning on lagged cultural activity, we are able to control for many unobserved components that lead individuals into cultural participation. Furthermore, since activity levels are the same within these two groups, we mitigate the problem that cultural participation influences the control variables, and also lagged health, via past engagement and unobserved confounders.



Source: Own illustration.

Figure 2.1: Matching within groups defined by past cultural participation

The matching analysis should be based on individuals in the region of common support or for which there is sufficient overlap between the treatment and control group. In other words, there must be a sufficient number of control individuals that have a relatively high probability of receiving the treatment. We impose the common support by dropping treated individuals whose propensity score is higher than the maximum or less than the minimum propensity score of the control individuals.

Furthermore, participants and non-participants should be similar in terms of their covariate distribution. This can be assessed by using a simple t-test to examine whether there are significant differences in covariate means between these groups. However, we will rely on the standardized bias in covariates to assess the covariate balance between treated and control units. According to Imbens (2015), the standardized bias is more robust to sample size than the t-test and should be preferred to assess the extent of covariate imbalance between the treatment and control group. The standardized bias is defined as the percentage difference in covariate means between treated and control individuals normalized by the standard deviation, and is calculated before and after matching (see also Rosenbaum and Rubin, 1985):

$$SB = 100 \cdot \frac{\overline{X_1} - \overline{X_0}}{\sqrt{(V_1(X) + V_0(X))/2}}, \tag{2.2}$$

where $\overline{X_1}$ ($\overline{X_0}$) is the mean of the covariate X in the treatment (control) group, and V_1 (V_0) is the respective variance. A remaining bias below 3 or 5 percent after matching is generally deemed to be sufficient (e.g. Caliendo and Kopeinig, 2008). As argued by Rosenbaum and Rubin (1985), any remaining standardized difference above 20 percent is considered as large.

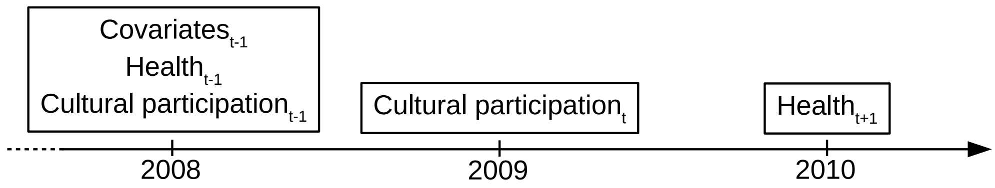
A final issue concerns the estimation of the variance of the parameter for the cultural participation variable. Generally, the estimated standard errors neglect the estimation of the propensity score. The uncertainty associated with the propensity-score estimation is thus disregarded, which could increase the variance of estimators (Heckman et al., 1998). We therefore follow Marcus (2014) and Schmitz and Westphal (2015) and use robust standard errors from the weighted least squares regressions. These standard errors are nevertheless compared with standard errors resulting from bootstrapping the regression-adjusted propensity-score matching procedure. Applying the bootstrap to calculate the standard errors is a very popular method in applied analyses, and Abadie and Imbens (2008) suggest that this approach might be valid in the case of the kernel matching estimator.

2.4 Data

2.4.1 Description of the Estimation Sample

This study uses longitudinal data from the Socio-Economic Panel (SOEP) study (Wagner et al., 2007).[2] It is a large and representative survey of German households which started in 1984. It is suitable for our purposes since it includes information on health status, demographic and socioeconomic background, leisure and social activities, personality and youth socialization. Figure (2.2) illustrates the basic structure of the estimation sample. The cultural participation variable is measured in 2009. While information on

[2]Socio-Economic Panel (SOEP), data for years 1984-2012, version 29, SOEP, 2013, doi:10.5684/soep.v29.



Source: Own illustration.

Figure 2.2: Structure of the estimation sample

cultural activities is available in other waves, the focus is on this year due to the abundant availability of both health measures and indicators for leisure and social activities around this year. The health outcomes are measured in 2010. Furthermore, we will use information on health outcomes and cultural participation in year 2008 to mitigate the problem of reverse causality and unobserved heterogeneity. Most of the control variables are gathered in 2008, to alleviate the problem that some background characteristics are influenced by the decision to visit cultural events or the anticipation of it.

The final analysis sample includes 4,158 individuals. The reasons for the relatively small number of cases are twofold: First, the estimation is based on individuals providing non-missing information on all (dependent and independent) variables. Second, we also include retrospective information on youth and childhood conditions potentially related to both cultural participation and health in adulthood. The questionnaire on youth and socialization was introduced in 2000, and has been completed by households that entered the SOEP henceforth. This leads to a further reduction in sample size.

2.4.2 Definition of Cultural Participation

In the 2009 wave, the SOEP survey includes a battery of questions directly related to the respondents' leisure and social activities. The respondents had to assess the frequency of various activities during their free time, such as doing sports, meeting with friends or political commitment. Cultural participation is measured based on the following item:

Going to cultural events (such as concerts, theater, lectures etc.)

The respondents had to check how often they do this activity on a four-point scale using the options "weekly", "monthly", "less often" or "never". Thus, we adopt a rather narrow definition of cultural participation that comprises both the performing arts and visual arts, which is however consistent with cultural economic approaches (e.g. Frey, 2008). The cultural attendance indicator distinguishes between those who often go to cultural events and those who rarely or never attend cultural events. The treatment group is confined to individuals who visit cultural events at least monthly (n = 923).

The control group comprises those respondents that less often or never visit cultural events ($n = 3,235$).

2.4.3 Measurement of Health Outcomes

As outcome variables we use generic measures of physical and psychological health that are available in the SOEP study since 2002. They are calculated based on the short-form 12 (SF-12) questionnaire which is a brief version of the SF-36 survey and a widely accepted and validated tool for the measurement of health-related quality of life in population surveys (Andersen et al., 2007). The SOEP version of the SF-12 consists of twelve self-reported items that comprehensively measure the respondents' physical and mental health. These items are merged into eight subscales and summarized into two aggregate dimensions via exploratory factor analysis: "physical health" (*pcs*) and "mental health" (*mcs*).[3] The *pcs* includes the subscales physical functioning, role physical, bodily pain, and general health perception. The *mcs* consists of the subscales mental health, role emotional, social functioning, and vitality. The main dimensions are standardized such that their mean equals 50 and their standard deviation equals 10. The individuals in our sample on average score slightly higher on the mental health scale (51.5) and slightly lower on the physical health scale (48.2) than the general population, but the difference seems rather negligible (see Table (2.A.1) in the appendix).

2.4.4 Definition of the Conditioning Set

The decision to attend cultural events is non-random and can be couched in terms of a constrained utility-maximization problem. Individuals maximize their utility by choosing the level of cultural goods consumption and other commodities under budget and time constraints. The demand for cultural goods, in turn, is a function of taste or preference for artistic and cultural experiences acquired in the past (e.g. Gray, 2011). We therefore include a large set of personal characteristics to capture the respondents' constraints and preferences with respect to the arts and health. The set of covariates can be differentiated into seven groups: demographic background, socioeconomic status, social activities, leisure activities, health behavior, personality traits, and early exposure to the arts (see Table (2.A.1) in the appendix).

As is standard in any empirical examination of cultural participation and health outcomes, we control for the respondent's sex, age, household size and urbanization level. The relationship between sex and age on the one hand and arts participation on the other hand is ambiguous (Seaman, 2006). The size of the family points to time and budget constraints that could shape the decision to attend cultural events, or to engage in other

[3]See Table (2.A.2) in the appendix for a detailed description of question wording and response scales in the SOEP.

artistic or health-related activities. The impact of the urbanization level possibly reflects the supply and accessibility of cultural facilities, which could influence the likelihood of attending cultural events (Gray, 2011).

To capture the respondent's socioeconomic status, we also include control variables for educational attainment (secondary, vocational, and tertiary), household income, and employment status (employed, not employed, unemployed). Previous empirical research has found a high correlation between educational level and arts attendance (Seaman, 2006). Consumption of cultural or arts goods requires investments in arts-specific human capital and tastes, to understand and appreciate artistic performances. Clearly, education could be a means to acquire theses skills. Furthermore, household income and employment status represent the financial and time resources necessary to visit cultural events. Finally, education, income and employment status have been shown to be highly correlated with health outcomes (for a review, see Cutler et al., 2011)

We include several markers for the respondent's social capital to control for the non-random selection into cultural participation due to socialization. Specifically, we take into account how often the individual volunteers, is politically engaged, goes to the church, and visits neighbors and relatives.

In contrast to previous studies, we include an extensive set of leisure activities. On the one hand, it seems plausible that those individuals that pursue an active lifestyle are more likely to attend cultural events and are generally healthier than less active persons. More important, we control for the extent of the individual's artistic activities. It could be argued that those persons who sing or play a musical instrument in their spare time are also more likely to attend cultural events. Exposure to a more focused form of creative activity could, for instance, reflect a general preference or taste for the arts (e.g. Gray, 2011). On the other hand, other leisure activities could reflect time constraints that reduce the opportunities to engage in both cultural and health enhancing activities.

Health behavior could also influence arts participation and perceived health. Therefore, we control for a variety of markers for the individual's health-related lifestyle, namely body-mass index (BMI), smoking status, alcohol intake and dietary behavior.

A further improvement compared to previous studies is the inclusion of personality traits among the covariates. These characteristics are usually unobserved in large-scale population studies, and potentially reflect systematic differences between individuals in terms of cultural participation and health. There is substantial evidence on the relationship between personality and health (Almlund et al., 2011). Moreover, personality traits can influence the decision to visit cultural events, and it has been noted that personality-related individual differences are critical for understanding arts preferences and appreciation (e.g. Kraaykamp and Eijck, 2005). For example, individuals with a general appreciation for arts are potentially more likely to derive satisfaction from artistic performances than other persons and hence more likely to attend cultural events. In the

SOEP questionnaire, the respondent's personality is assessed with the Big Five personality inventory. Personality differences can thus be traced back to five main personality traits: neuroticism, extraversion, openness to experience, agreeableness, and conscientiousness (Richter et al., 2013, pp. 44). The Big Five personality traits have been included in the 2005 and 2009 waves of the SOEP, and we use the average of both years as control variables. This is a valid approach assuming that the respondent's personality traits are rather stable over the life cycle (see also Costa and McCrae, 1988).

From an economic perspective, current arts participation is influenced by past cultural activities (Throsby, 1994). Cultural economists stress that adult demand for cultural engagement is heavily influenced by childhood experiences (e.g. Gray, 1998; Morrison and West, 1986). To approximate early exposure to the arts, we use retrospective information on the respondent's socialization in childhood and youth. Specifically, we include the educational level of the parents, the place the respondent lived during childhood, whether the individual did sports and attended musical lessons during youth, and the number of siblings.

2.5 Results

2.5.1 Linear Regression Estimates

We first present results from simple linear regression models of the outcomes, where the set of covariates is successively included. The first column of Table (2.1) reports the unconditional (raw) association between cultural participation on the one hand and the physical and mental health summary scores on the other hand. It shows that those individuals who frequently visit cultural events are on average healthier than less culturally active persons. A discrete change in the cultural attendance indicator from zero to one increases the physical and mental health scores by 1.4 and 2.8 points (or 14 and 28 percent of a standard deviation), respectively. This relationship is significant at the 99 percent level.

In general, including the covariate groups gradually decreases the positive relationship between cultural attendance and health, while the standard errors remain comparatively stable across the different specifications. This suggests that the (raw) correlation between arts participation and health reflects positive selection into cultural event visits based on observable characteristics. The second column of Table (2.1), however, shows that conditioning on the respondent's demographic background (i.e. sex, age, household size, family status and place) increases the effect on physical health, but lowers the impact on mental health. The different patterns are possibly age-related. On the one hand, spectators of arts performances tend to be older but at the same time have more physical health problems, resulting in an underestimate of the physical health effects of arts participation.

Table 2.1: Simple linear regression results

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Health outcomes								
Physical health	1.379***	2.491***	1.237***	1.096***	0.845**	0.642*	0.546	0.549
	(0.357)	(0.327)	(0.334)	(0.341)	(0.357)	(0.353)	(0.345)	(0.346)
Mental health	2.765***	2.158***	1.662***	1.354***	1.148***	1.065***	0.732**	0.755**
	(0.325)	(0.330)	(0.356)	(0.362)	(0.377)	(0.377)	(0.353)	(0.353)
Control variables								
Demographics		X	X	X	X	X	X	X
Socioeconomic Status			X	X	X	X	X	X
Social activities				X	X	X	X	X
Leisure activities					X	X	X	X
Health behavior						X	X	X
Personality traits							X	X
Early exposure								X

Note: Number of individuals=4,158. Robust standard errors, clustered at the individual level, in parentheses. The *health outcomes* are based on the SF-12 health survey and measured in 2010. *Cultural participation* is assessed in 2009 and measures at least monthly cultural activitiy. The control variables are measured in 2008, except for personality traits and early exposure to the arts. Personality traits are calculated based on the average scores in 2005 and 2009. Early exposure to the arts is based on retrospective information regarding childhood and youth. Significance levels: *** $p<0.01$, ** $p<0.05$, * $p<0.1$.

Source: Own calculations based on SOEP v29.

On the other hand, older individuals tend to be happier than younger persons and thus exhibit less mental health problems. This would imply an overestimate of the mental health benefits of cultural participation.

With respect to the remaining control variables, the physical and mental health scores follow similar patterns. As shown by the third column, conditioning on socioeconomic status (i.e. education, income, and employment status) mitigates the potential health benefits of arts attendance. Clearly, better-educated or higher-income individuals are possibly more likely to visit cultural events and have less health problems, because they have the financial resources and cognitive competencies to afford cultural goods and deal with or prevent diseases. Hence, ignoring the selection based on socioeconomic status would bias our estimates upwards.

The coefficients reported in the fourth column suggest that part of the arts-health relationship can be explained by social activities. As argued above, cultural activities most likely involve social interactions that could benefit health. What is more, individuals who socialize much or are politically active tend to visit cultural events more often. This indicates that individual-level social capital is an important omitted or unmeasured variable in previous studies which explains at least part of the health effects of arts participation. However, the problem of selection due to social capital seems to be less pronounced than that of socioeconomic status.

Including leisure activities and health behaviors in the fifth and sixth column further weakens the link between arts attendance on health. It appears that those individuals who generally follow an active and healthy lifestyle are more likely to visit cultural events and have less health problems. Thus, not controlling for the level of leisure and health-related activities may overstate the health benefits of arts participation.

The seventh column of Table (2.1) shows that there is some selection based on personality traits. On the one hand, open-minded individuals are possibly more likely to attend cultural events and exhibit less health problems. On the other hand, respondents with high scores in neuroticism or low scores in emotional stability are probably less likely to engage in cultural participation and have worse health outcomes. Hence, positive selection based on favorable personality traits may qualify previous estimates as upward biased.

The last column reports estimates conditioning on all covariates including the indicators for childhood exposure to the arts. However, the coefficients for the binary cultural attendance variable basically remain unchanged, and it seems that childhood experience is unrelated to the arts-health relationship in our sample. On the one hand, this may be attributed to recall bias where individuals misjudge, for example, their parents' education or the extent of physical and artistic activity during childhood and early adolescence. To the extent that this type of error is systematic, the effect of childhood conditions on adult health and cultural participation might cancel out. On the other hand, Germany has experienced an expansion of educational opportunities in the past 50 years. This implies that children who were born just before or during this period tend to have better educational outcomes than their parents. These children might have been able to "compensate" for their parents' low educational level or socioeconomic status, and display health outcomes and participation rates in adulthood similar to children from better-off families. Hence, it appears plausible that childhood conditions are rather unrelated to adult health and cultural activities in Germany.

2.5.2 Matching Estimates

Propensity Score Model

The first step of the matching analysis involves the estimation of the respondents' propensity scores. The individual-specific predicted probabilities from this regression are then used as the propensity scores in the matching procedure. The results provide interesting insights into the correlates of cultural participation in our sample. As illustrated in Figure (2.A.1) in the appendix, age positively influences the probability of regular cultural participation, adjusting for the remaining covariates. The propensity score rather continuously increases with age, and the highest participation rates are observed among individuals aged 65 and older. This could reflect the greater availability of time for leisure and social activities after retirement.

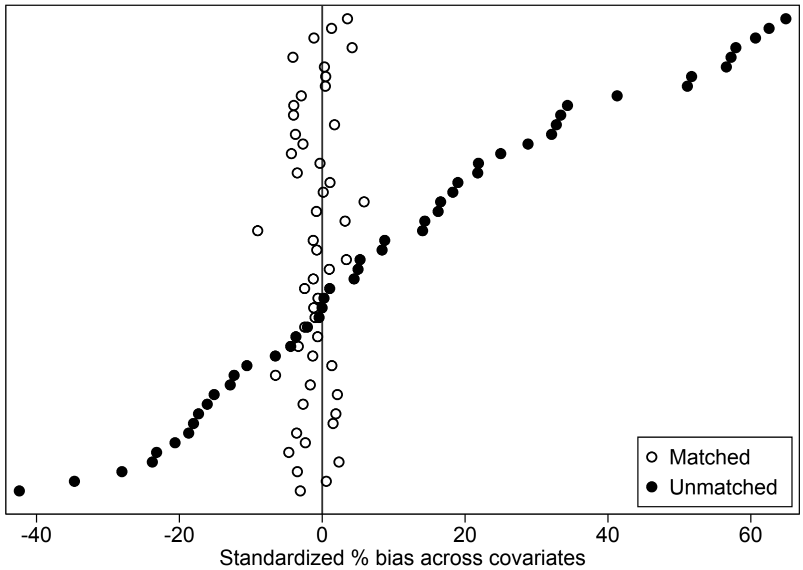
The estimates for the remaining covariates are shown in Table (2.A.3) in the appendix. Higher propensity scores are observed for individuals who are single or divorced. Better education seems to increase cultural participation rates. Individuals with an intermediate or academic school degree have a higher probability of visiting cultural events than those with a basic education. Having a vocational degree, in contrast, is associated with less frequent cultural event visits. Furthermore, arts participation increases with the logarithm of household income and is positively related to unemployment. This finding might seem somewhat surprising since unemployment is usually associated with loss of income. However, the majority of persons that were given supplementary questions with respect to socialization in youth in the SOEP belongs to samples that were included in more recent years. High-income households might therefore be overrepresented in our sample, since they were included only in 2002. Hence, the income loss due to unemployment for individuals living in these households might be less severe and the positive effect of unemployment on cultural participation likely reflects more time available for leisure activities.

Culturally active persons also tend socialize more and exhibit higher levels of political and civic engagement. What is more, cultural attendance is positively correlated with artistic activities in the leisure time. Moreover, individuals who smoke less, follow a health-conscious diet, and have lower BMI scores are also more likely to visit cultural events often. As expected, individuals that score high on the openness (neuroticism) trait have a higher (lower) propensity of attending arts activities. Finally, childhood exposure to the arts, approximated by musical activity in youth, parental education, place and number of siblings seems to be unrelated to adult consumption of art performances.

Covariate Balance

After calculating the propensity scores, the matching procedure as outlined in Section (2.3) is employed. The aim is to find control individuals that are similar to cultural participants in terms of their covariate distribution. Hence, it is a critical task in any matching analysis to assess the covariate balance between these two groups. This is usually done by evaluating the standardized difference in each independent variable. We will do this for each covariate.

To illustrate this step, Table (2.A.4) in the appendix reports mean values and the standardized bias for each covariate among participants and non-participants, and before and after matching. The numerical results are illustrated in Figure (2.3). It graphically shows the standardized percentage bias between cultural participants and non-participants for each covariate and before (solid circles) and after (hollow circles) match-



Source: Own calculations based on SOEP v29.

Figure 2.3: Standardized covariate bias before and after matching

ing. Generally, the matching algorithm performs well in terms of bias reduction. The normalized differences are considerably lower after matching and are less than or close to 5 percent. One exception is the variable which indicates whether the respondent's parents have a university degree. The standardized bias for this variable amounts to 9 percent after matching. This, however, implies a significant reduction of covariate imbalance compared to the non-matched sample, and is still close to the arbitrarily defined thresholds mentioned by Caliendo and Kopeinig (2008) (3 or 5 percent). We will complement the matching procedure with additional regression adjustment, which should reduce any bias emanating from remaining covariate imbalances.

A comparison of aggregate sample statistics before and after matching supports the conclusion that the overall matching quality is satisfying. The pseudo-R^2 figures in Table (2.2) emanate from a regression of the propensity score on the covariates using the unmatched (raw) and matched sample, respectively. These quantities suggest that the explanatory power of the regressors is fairly low in the matched sample ($R^2 = 0.01$) compared to the unmatched sample ($R^2 = 0.3$). This was to be expected as there should be no systematic differences in covariate distributions between treated and control individuals after matching (Caliendo and Kopeinig, 2008). Furthermore, a likelihood ratio test for the hypothesis that all coefficients are zero is rejected before matching, but cannot be rejected after matching. This is again in accordance with the expectation that the propensity to visit cultural events is unrelated to observable characteristics after matching. Finally, the mean standardized difference after matching is equal to 2.1 while the median bias amounts to 1.5, reflecting a considerable reduction in terms of covariate imbalance compared to the unmatched sample.

Table 2.2: Overall statistics on covariate balance

Sample	Pseudo R^2	LR χ^2	$p > \chi^2$	Mean %Bias	Median %Bias
Unmatched	0.293	1290.57	0.000	13.6	7.8
Matched	0.012	30.35	1.000	2.1	1.5

Source: Own calculation based on SOEP v29.

Main Results

Table (2.3) provides regression estimates using the matched sample of culturally active and inactive individuals. They basically emanate from a weighted linear regression of the outcome on the cultural participation indicator and covariates, where the weights are calculated based on the propensity-score matching procedure described above. The figures in parentheses are robust standard errors from these weighted linear regressions.

Table 2.3: Matching results

	(1)	(2)	(3)
Physical health			
Coef.	0.394	0.063	0.181
	(0.374)	(0.304)	(0.296)
$N_{\text{Participants}}$	899	898	893
N_{Controls}	3235	3235	3235
$N_{\text{Off support}}$	24	25	30
Mental health			
Coef.	0.898**	0.731**	0.631*
	(0.408)	(0.367)	(0.354)
$N_{\text{Participants}}$	899	901	894
N_{Controls}	3235	3235	3235
$N_{\text{Off support}}$	24	22	29
Control variables	Yes	Yes	Yes
Lagged health	No	Yes	Yes
Unobserved heterogeneity	No	No	Yes

Note: Robust standard errors from weighted linear regression in parentheses. The *health outcomes* are based on the SF-12 health survey and measured in 2010. *Cultural participation* is assessed in 2009 and measures at least monthly cultural activitiy. The control variables are measured in 2008, except for personality traits and early exposure to the arts. Personality traits are calculated based on the average scores in 2005 and 2009. Early exposure to the arts is based on retrospective information regarding childhood and youth. *Lagged health* refers to health status in 2008. In the third column the coefficients represent the average of matching estimates based on matching within groups defined by the same level of past cultural activitiy in 2008. Significance levels: *** p<0.01, ** p<0.05, * p<0.1.

Source: Own calculations based on SOEP v29.

The first column of Table (2.3) displays the estimates from the matching model controlling for all observable covariates. For physical health it is equal to 0.394, which is smaller than the OLS estimate but still insignificant. The parameter for mental health equals 0.898, which is significant at the 95 percent level and higher than the OLS result. The difference between the matching and linear regression results can be explained by the different weighting schemes they apply (e.g. Angrist and Pischke, 2008). Hence, using the matched sample of cultural participants and non-participants, we find that regular consumption of arts activities increases the mental health index by roughly 9 percent of a standard deviation.

The second column of Table (2.3) reports the estimated associations between cultural attendance and health when we control for reverse causality. Specifically, we include the respective health outcome measured in 2008 among the covariates. As a result, the point estimates in both the physical and mental health equations decrease. This provides evidence for a non-random selection of healthy individuals into cultural attendance. However, the coefficients are more precisely estimated and the mental health benefits remain statistically significant.

In the third column, we take unobserved heterogeneity into account by providing weighted average estimates from matching analyses performed in the two strata defined by past cultural participation. Since baseline cultural activitiy is the same in these two groups, this approach reduces the impact of unobserved variables on the selection process. The coefficient in the physical health regression remains insignificant. The point estimate in the mental health equation decreases further but remains significant at the 10 percent level. Hence, it is possible that there is residual bias due to unobserved confounders positively influencing arts participation and psychological well-being.

2.6 Robustness Checks and Sensitivity Analysis

2.6.1 Variation of the Matching Procedure and Parameters

This section provides results regarding the robustness of our estimates with respect to variations of the matching procedure and parameters. The focus is on the matching model taking reverse causality and unobserved heterogeneity into account. In the following we restrict the analysis to mental health, given the lack of significant results for physical health. The findings are displayed in Table (2.4). In column 1, we repeat the baseline matching estimate for mental health. In column 2, we re-estimate our preferred specification using a bandwidth equal to 0.03 instead of 0.06. The bandwidth choice generally involves a trade-off between the bias and variance of point estimates (see also Caliendo and Kopeinig, 2008). A larger bandwidth puts greater weight on control individuals with more distant propensity scores, elevating the risk of using poor matches. One the one hand, this could increase the bias while, on the other hand, variance is reduced since more observations are used in the calculation of the ATT. Choosing a smaller bandwidth, to the contrary, reduces bias and increases the variance because it puts greater weight on

Table 2.4: Robustness analysis of the matching procedure (mental health only)

	(1) Baseline	(2) Bw=0.03	(3) NN	(4) Sig. controls	(5) Exog. controls	(6) Thick support	(7) Boot. SE
Coef.	0.631*	0.646*	0.836**	0.618	1.055***	1.367*	0.631
	(0.354)	(0.357)	(0.389)	(0.388)	(0.366)	(0.827)	(0.396)
$N_{\text{Participants}}$	894	893	894	909	908	236	894
N_{Controls}	3235	3235	3235	3235	3235	251	3235
$N_{\text{Off support}}$	29	30	29	14	15	9	29
Control variables	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Lagged health	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Unobserved heterogeneity	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Note: Robust standard errors from weighted linear regressions in parentheses. The *health outcomes* are based on the SF-12 health survey and measured in 2010. *Cultural participation* is assessed in 2009 and measures at least monthly cultural activitiy. The control variables are measured in 2008, except for personality traits and early exposure to the arts. Personality traits are calculated based on the average scores in 2005 and 2009. Early exposure to the arts is based on retrospective information regarding childhood and youth. *Lagged health* refers to health status in 2008. The coefficient in each column represents the average of matching estimates based on matching within groups defined by the same level of past cultural activitiy in 2008. Significance levels: *** p<0.01, ** p<0.05, * p<0.1.

Source: Own calculations based on SOEP v29.

similar matches, but uses fewer comparison units from the control group. The smaller bandwidth reduces the number of treated individuals in the region of common support only slightly. However, narrowing the bandwidth produces point estimates and standard errors that are almost identical to the results obtained with the larger bandwidth.

Column 3 shows the results from a propensity-score matching method using a variant of nearest-neighbor matching (see, for example, Marcus, 2014; Morgan and Winship, 2007). The basic idea of nearest-neighbor matching is to find a match for each participant that is closest in terms of the propensity score. That is, instead of using the weighted average of all individuals in the control group, the counterfactual is calculated based on only one or several comparison individuals. If only one nearest neighbor is used, the matched individual receives a weight equal to 1. In case several nearest neighbors are used, the weight for matched individuals equals 1 divided by the maximum number of nearest neighbors. The weight for unmatched control individuals is equal to zero. The number of nearest neighbors also involves a trade-off. Using more comparison units or information to calculate the counterfactual clearly reduces the variance of the effect estimates, but it could increase the bias due to the inclusion of poor matches.

We employ nearest-neighbor matching with replacement using 5 nearest neighbors. Thus, an individual from the comparison group can be used multiple times for the calculation of the counterfactual of each cultural participant. Furthermore, to minimize the problem of using poor matches, we impose a caliper equal to 0.25 standard deviations of the estimated propensity score (Rosenbaum and Rubin, 1985). The caliper represents the maximum propensity-score distance between treatment and control cases. The control individual must lie within this predetermined range to be included in the calculation of the counterfactual outcome. This should presumably improve the quality of the matched control subjects but could also increase the estimated variance, since fewer information is used to calculate the counterfactual. As shown in column 3, the nearest-neighbor procedure generates slightly larger point estimates and standard errors, which could be attributed to the loss of information when the outcome of the comparison group is calculated based on the five control individuals with the nearest propensity score.

In column 4, we use an alternative set of conditioning variables. The choice of the control variables is based on previous empirical work and theoretical considerations. However, there is a dispute on which and how many variables should be included in propensity-score matching analyses. According to, for example, Caliendo and Kopeinig (2008) there are two pitfalls associated with the inclusion of too many covariates: First, including irrelevant variables could make it difficult to find matches for treated individuals, reducing the area of common support. Second, the inclusion of nonsignificant variables could increase the variance of the propensity score estimates. We follow the strategy proposed by Marcus (2014) and employ a forward-selection search for the propensity-score model. That is, we estimate a probit model with cultural participation as the dependent variable and successively add the covariates. A covariate is kept when it is significant at the 5 percent level. As shown in column 4, the point estimate conditional on the subset of significants covariates is slightly smaller but comparable to the baseline estimate.

A further issue relates to the question whether some of the control variables are endogenous, that is whether they are influenced by cultural attendance. As argued by, for example, Rosenbaum (1984), the conditioning set should only include variables that are unaffected by the treatment. Although measured one period before, some covariates can be affected by the anticipation of the treatment. Individuals might already know whether they attend a cultural event in the future, which might change behavior in the current period (for a similar reasoning, see for example Lechner, 2008). This is particularly true for social and leisure activities that are strongly correlated with cultural participation. In this case, there is obviously a trade-off between reducing the bias due to omitted variables and bias resulting from endogenous covariates.[4] Column 5 shows what happens to the estimates when we exclude potentially endogenous variables, and condition on presumably exogenous covariates only. These are the variables that are either time-invariant or cannot be easily adjusted to the anticipation of the treatment.[5] The association between cultural participation and mental health clearly increases. Thus, reducing the bias due to endogenous control characteristics might come at the cost of increasing bias attributable to unobserved confounders.

[4] However, the bias induced by endogenous control variables is negligible if the unconfoundedness assumption holds (Lechner, 2008).

[5] These include: sex, age, household size and composition, family status, agglomeration level, education, household income, employment status, personality traits, musical and sports activities during youth, parental education, place during childhood, number of siblings, and health-related lifestyle.

Column 6 shows the matching estimates using a narrower range of propensity scores. The region of common support in the subsamples defined by past cultural participation ranges from very low to very high propensity-score values (0.0002 to 0.907 for non-active individuals, and 0.022 to 0.989 for active individuals). As argued by Black and Smith (2004), matching estimates based on a wide propensity-score distribution entail a variety of problems: First, the effects of unobserved heterogeneity may amplify when matching relies on very high or very low values of the propensity score. Second, one might also be concerned about individuals with high propensity scores who do not participate in cultural activities. This could reflect measurement error in the cultural participation variable among these respondents. Third, using the whole region of common support might mask heterogeneous effects of cultural participation across different propensity scores and individuals, respectively.

Black and Smith (2004) suggest to focus on propensity-score values between 0.33 and 0.67, which they call the region of "thick support". Column 6 shows that this greatly increases the cultural attendance coefficient. The higher point estimate could simply reflect the mean effect of cultural engagement for a narrower population defined by the region of "thick support". However, it also points to the existence of measurement error in the cultural participation indicator.

In our context, we do not know the true participation probability and the propensity score has to be estimated. This poses a serious problem to the variance of our matching estimator. It has been shown by Heckman et al. (1998) that the variance due to the estimation of the propensity score adds to the variance of average treatment effects. Our matching estimator does not take the uncertainty associated with the estimation of the propensity score into account. Instead, we rely on robust standard errors from the weighted regressions. To further assess the robustness of the results, we also computed bootstrapped standard errors. The bootstrap procedure involves the following steps: (*i*) Draw a random sample with replacement from the observed sample, (*ii*) estimate the propensity score, (*iii*) compute the weights for matched individuals, (*iv*) perform weighted regressions using robust standard errors. The bootstrap is repeated 1,999 times and the bootstrap standard error is obtained by calculating the standard deviation of the bootstrapped parameter estimates according to the following formula (see also MacKinnon, 2006):

$$s^*(\hat{\beta}) = \sqrt{\frac{1}{B-1}\sum_{b=1}^{B}(\hat{\beta}_b^* - \bar{\beta}^*)^2},$$

where B is the number of bootstrap replications, $\hat{\beta}$ is the original parameter estimate, $\hat{\beta}_b^*$ is the corresponding estimate for the bth bootstrap replication, and $\bar{\beta}^*$ is the mean of the $\hat{\beta}_b^*$. According to Abadie and Imbens (2008), bootstrapping could be valid in the case of kernel matching which is asymptotically linear and with which the number of matches increases in the sample size. As shown in column 7, bootstrapped standard error is slightly more conservative than the robust standard errors obtained from the weighted linear regression.

To sum up, the estimated association between engagement in cultural activities and mental health is rather robust to the application of different matching algorithms and sets of conditioning variables. However, the estimate based on the region of thick support suggest that the effect of cultural participation might vary depending on the (sub)population considered. Moreover, measurement error in the cultural attendance variable might bias the results. The next section provides a more thorough assessment of the estimation results to more or less plausible deviations from unconfoundedness.

2.6.2 Sensitivity Analysis of Point Estimates

The estimated significant and positive relationship between cultural attendance and mental health may occur due to unobserved variables that are simultaneously related to engagement in cultural activities and psychological well-being. Obvious candidates in our context are skills and preferences regarding the consumption of arts goods and services. If unobserved components of abilities and taste positively influence both cultural participation and mental well-being, our estimates likely overestimate the benefits of arts participation.

To check the sensitivity of the matching estimates to unobserved confounding, we rely on a method proposed by Ichino et al. (2008). Generally, their approach allows the researcher to assess the responsiveness of matching estimates depending on various hypothesized deviations from unconfoundedness. Starting point is an unobserved binary variable U. They assume that including this variable among the matching covariates would completely eliminate the bias resulting from unobserved factors. The values of U for active and inactive individuals are calculated on the basis of four parameters $p_{ij} = Pr(U = 1|T = i, Y = j)$, with $i, j \in \{0, 1\}$, where T and Y are indicators for arts participation and health scores above the sample median[6], respectively. Thus, the distribution of the unobserved binary variable can be characterized by the probability that $U = 1$, depending on the membership of the individual in each of the four groups defined by participation status and the outcome. Once the value of U is calculated for each individual, one can include it as an additional matching variable and estimate the relationship between arts and health conditional on observed covariates and the unobserved binary confounder. The simulation of U and the estimation step are repeated 200

[6]As proposed by Nannicini et al. (2007), the continuous health index is dichotomized because the analysis framework applies to binary outcomes. We use the sample median as a cutoff point. The relationship between cultural participation and health is, nevertheless, estimated for the continuous outcome.

times with the respective parameters p_{ij}, that is we perform 200 random draws from a uniform distribution.[7] Thus, at each iteration the values of U (0 or 1) are randomly allocated across individuals given the fixed empirical distribution of U according to the parameters p_{ij}. By averaging the results over all iterations, we obtain an estimate for the health benefit of cultural participation which is presumably unbiased under pre-specified hypotheses concerning unconfoundedness.

It is important for this type of sensitivity analysis to find "reasonable" values for the parameters p_{ij}, inevitably leading to the question which type of unobserved confounder could invalidate our findings (Nannicini et al., 2007). The greatest threat to our estimates comes from unobserved factors that increase mental health even in the absence of cultural participation and that lead individuals into cultural attendance. Suppose, for instance, that the unobserved factor U represents the individual's taste or preference for the arts. Innate or acquired taste increases the probability of attending cultural events. At the same time, it might also improve health in the absence of cultural participation. The effect of taste on participation is called *selection effect* and can be defined as $s = p_{1.} - p_{0.}$. The effect of taste on the potential outcome among non-active individuals can be characterized as the *outcome effect* and is calculated as $d = p_{01} - p_{00}$. Positive selection effects thus exist if cultural participants are more likely to have a taste for the arts than non-participants. A positive outcome effect means that taste for the arts is more common among healthy persons, therefore improving the health of non-participants under the counterfactual scenario. However, d and s only measure the sign of the influence of U. To gain insights about the size of associations, we additionally estimate at each iteration a logit model of both the counterfactual health outcome and the cultural participation indicator, including U and observed characteristics as predictors. The average odds ratio of U measures the outcome effect Γ on the one hand, and the selection effect Λ on the other hand. There is obviously a close relationship between d and s on the one hand and Γ and Λ on the other hand. More importantly, providing the magnitude of the outcome and selection effect as an additional output sheds light on the plausibility of the hypothesized deviation from unconfoundedness. For example, a significant association between arts attendance and health may disappear when adjusting for the unobserved variable, but only for unreasonably high outcome and selection effects.

Ichino et al. (2008) propose two methods to determine the values of values of p_{ij}, and hence d and s. First, we impute p_{ij} by simulating a confounder U that emulates observed covariates. Thus, the distribution of U is equal to the empirical distribution of observable characteristics. In order to include all factors, we use binary transformations of the continuous variables that indicate values above the respective sample mean. Using this approach, we can examine the robustness of our estimates to the inclusion of an

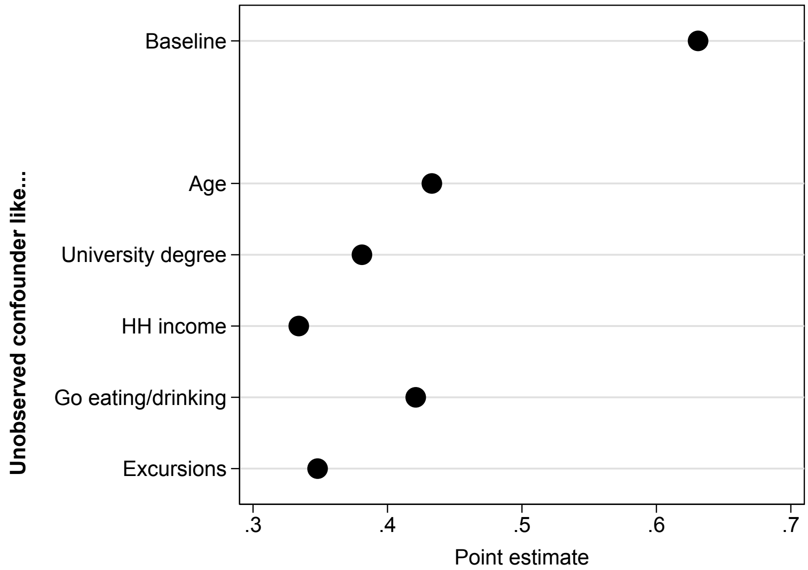
[7] We use a slightly altered version of the Stata do file *sensatt* developed by Nannicini et al. (2007) to conduct the sensitivity analysis.

unobserved confounder as a control variable that is similar to the observed covariate used to simulate the distribution of U. Second, we perform a rather exploratory sensitivity analysis and search for values of the parameters p_{ij} that could, if we control for the respective U, render the health benefits of arts participation insignificant or zero. The plausibility of such a finding can be assessed by examining the magnitude of the outcome and selection effects under this scenario. The values p_{ij} can be derived by solving a system of equations for desired levels of d and s.[8] We end up with a set of parameters p_{ij} that underly a specific combination of $d > 0$ and $s > 0$. We estimate the association between arts participation and health including observed covariates and the variable U for different combinations of d and s ranging from 0.1 to 0.5.

The sensitivity analysis is conducted for mental health, since it only makes sense for significant findings. The full results of the first simulation exercise are shown in Table (2.A.5) in the appendix. It shows how the baseline estimate changes when U emulates observed covariates and is included as an additional conditioning variable. The table provides information on the parameters d, Γ, s, and Λ which measure the direction and magnitude of the outcome and selection effect, respectively. The simulation results suggest that confounding due to an unobserved variable seems to be more frequent and severe if it leads individuals into cultural engagement and improves mental health in the absence of engagement ($s > 0$, $d > 0$). As an illustration, Figure (2.4) compares the baseline estimate assuming no unobserved confounding with selected point estimates from different estimations including the simulated confounder U among the covariates. We present the coefficients when U mimics the empirical distribution of age, university degree, household income, go eating/drinking or excursions. Unobserved confounding appears to be largest when the unobserved factor behaves like these variables. Generally, the coefficient for the cultural participation variable decreases substantially. For instance, if U has the same selection and outcome effects as *Excursions*, the effect of cultural attendance on mental health reduces from 0.63 to 0.34, which is roughly equal to a 46 percent decline.

One could argue that these findings just reflect the fact that U mimics the behavior of observable characteristics with positive outcome and relatively strong selection effects. To learn something about how strong the outcomes and selection effects have to be to drive the baseline estimate down to zero, we perform the simulation exercise for different combinations of $d > 0$ and $s > 0$. Table (2.5) shows that for increasing values of d and s the estimated mental health gain decreases. Low values of the outcome and selection effect, that is $d = 0.1$ and $s = 0.1$, already reduce the point estimate for the cultural attendance coefficient. Given s, higher values of d reduce the point estimate. However, values of $d > 0.1$ imply outcome effects equal to $\Gamma > 2.3$, which is larger than the maximum of outcome effects among observed covariates (see Table (2.A.5)). We therefore focus on

[8] See the section "The Derivation of the Sensitivity Analysis Parameters" in the appendix.



Source: Own calculations based on SOEP v29.

Figure 2.4: Baseline coefficient vs. selected point estimates under different assumptions about unobserved confounding

the more plausible case that $d = 0.1$ and $\Gamma \in [1.5, 1.6]$. Thus, we provide an informal test in the style of Altonji et al. (2005), by assuming that unobserved confounding being larger than selection on observables is unlikely. For a value of $s = 0.3$, given $d = 0.1$, the coefficient halves compared to the baseline estimate. Thus, when the unobserved variable is allowed to have an effect on outcomes and selection similar to observable covariates, the positive impact of cultural participation could diminish substantially. Larger selection effects may render the mental health gain zero, but only for implausibly high influences on the selection process.

The simulation analysis suggest that our estimates of the relationship between cultural participation and mental health are not robust to certain plausible deviations from the unconfoundedness assumption. Although controlling for a variety of observable characteristics that influence the probability of arts participation and mental well-being, and conditioning on past cultural activity, we are not able to fully take the selection process into account. Unobserved components of taste and skills related to the arts which heavily influence selection into cultural attendance could still explain why arts participants are healthier than non-participants.

2.7 Conclusion

This study examined the relationship between cultural participation and self-rated physical and mental health. Using rich individual-level data from the SOEP study, we found that the association of cultural engagement with health is heavily influenced by the im-

Table 2.5: Sensitivity analysis: unobserved confounder with positive selection and outcome effects (mental health only)

	$s = 0.1$ $\Lambda \in$ $[1.3, 1.5]$	$s = 0.2$ $\Lambda \in$ $[2.1, 2.3]$	$s = 0.3$ $\Lambda \in$ $[3.4, 3.7]$	$s = 0.4$ $\Lambda \in$ $[6.1, 6.6]$	$s = 0.5$ $\Lambda \in$ $[13.1, 14.0]$
$d = 0.1$, $\Gamma \in [1.5, 1.6]$	0.523	0.429	0.304	0.188	0.052
	(0.364)	(0.375)	(0.402)	(0.426)	(0.459)
$d = 0.2$, $\Gamma \in [2.3, 2.4]$	0.446	0.253	0.067	-0.141	-0.425
	(0.365)	(0.376)	(0.398)	(0.416)	(0.455)
$d = 0.3$, $\Gamma \in [3.6, 3.9]$	0.385	0.098	-0.192	-0.515	-0.907
	(0.369)	(0.374)	(0.392)	(0.413)	(0.448)
$d = 0.4$, $\Gamma \in [5.7, 6.6]$	0.342	-0.051	-0.430	-0.872	-1.387
	(0.372)	(0.378)	(0.392)	(0.411)	(0.436)
$d = 0.5$, $\Gamma \in [10.0, 11.7]$	0.288	-0.204	-0.689	-1.261	-1.857
	(0.381)	(0.377)	(0.389)	(0.413)	(0.430)

Note: Number of individuals=4,158. Robust standard errors from weighted linear regressions in parentheses. d measures the difference in the incidence of the unobserved variable U between healthy and unhealthy individuals in the absence of cultural participation ($p_{01} - p_{00}$). s measures the difference in the incidence of U between participants and non-participants ($p_{1.} - p_{0.}$). The full set of distribution parameters p_{ij} is obtained by solving a system of equations as illustrated in the appendix. Γ and Λ are odds ratios obtained from a logit regression including U and all covariates as independent variables, and quantify the effects of U on the untreated outcome and participation, respectively.

Source: Own calculations based on SOEP v29.

pact of personal and family background characteristics on selection and health outcomes. In addition to standard demographic and socioeconomic indicators, we also included social and leisure activities, healthy-related lifestyle, personality traits, and early influences on arts participation. Linear regression models including the full set of covariates only found a positive and significant correlation of regular arts attendance (at least monthly) with mental health.

We also adopted a propensity-score matching approach to reduce observed differences between cultural participants and non-participants and to create a sample of cultural participants and non-participants that approximately have the same observable characteristics. The problem of reverse causality and unobserved heterogeneity was addressed by exploiting the longitudinal nature of the dataset. When the lagged health outcomes were included as matching variables, the effect of arts participation on health decreased. This suggests that healthy individuals systematically select into cultural attendance. Concerns about unobserved confounding were reduced by performing matching within strata defined by past cultural activity. We found that this further reduces the mental health benefit of regular cultural engagement, but the point estimate remained significant. This finding was generally robust to different matching algorithms, but may depend on the specification of the conditioning set and the region of common support. The instability of point estimates may arise from heterogeneous cultural participation effects across individuals, or from measurement error in the arts participation variable.

Our empirical approach, however, did not completely solve the endogeneity problem. The greatest threat to significant health effects comes from unobserved factors that simultaneously influence the probability of regular arts participation and health outcomes. Sensitivity analyses indicated that plausible deviations from the unconfoundedness assumption were sufficient to substantially reduce the effect of engagement in arts activities on mental well-being. To sum up, this study was not able to find a robust effect of cultural participation on health despite including a rich set of covariates and addressing issues related to the selection process and unobserved heterogeneity. Unobserved components of taste and skills related to the arts may still create upward-biased estimates.

Thus, our study casts doubt on the effectiveness of cultural policy to reduce health inequalities. Despite widespread public funding of arts and cultural organizations, arts participation is still more prevalent among higher social status or education groups, also in Germany (Authoring Group Educational Reporting, 2012). One explanation could be that the supply of arts performances predominantly reaches those who have acquired arts-related preferences and abilities, for instance due to early exposure to the arts and parental influences, or frequent consumption of arts activities. Cultural policy should therefore involve means to provide financial incentives for arts participation and to alter preferences among lower social status groups. An effective way to stimulate cultural attendance among disadvantaged individuals could be the distribution of cultural vouchers. According to Frey (2008), vouchers can be an effective way to stimulate the demand for cultural events, because it loosens budget restrictions by lowering the price of cultural attendance. Furthermore, exposure to the "costless" arts experience may create a positive shock in preferences related to the arts. The combination of both effects can stimulate cultural participation in the future even in the absence of the free ticket.

Clearly, to answer the question whether cultural attendance actually affects health and which cultural policy instruments are effective is a matter of future research. This study could not identify true causal effects of regular arts consumption, but it adopted a research design which comes closer to a causal analysis than previous observational studies. The results presented in this study can thus be seen as a starting point for further investigations into the causal effects of cultural participation. By using more elaborate methods — such as panel-data analysis, instrumental variables, or randomized experiments — future studies can disentangle causal effects of arts participation from unobserved confounding. It is crucial from a policy perspective to obtain consistent evidence on whether visiting cultural events actually improves health, or whether the observed correlations reflect positive selection due to unobserved traits.

Appendix

Table 2.A.1: Summary of variables

Variable	Description	Mean	SD	Min.	Max
Treatment variable					
Cultural attendance	1=at least monthly, 0=less often or never	0.22	0.42	0	1
Health outcomes					
Physical health	Physical component summary scale (pcs)	48.20	10.02	15.02	73.08
Mental health	Mental component summary scale (mcs)	51.48	9.80	7.65	76.54
Demographic characteristics					
Female	1=female, 0=male	0.52	0.50	0	1
Age	Age of the individual (dummy variables)	55.61	14.64	24	99
Household size	Log. number of persons in hh	0.81	0.47	0	1.95
Number of children	Log. number of children (age<16) in hh	0.23	0.42	0	1.79
Married	1=married, 0=otherwise	0.69	0.46	0	1
Separated	1=separated, 0=otherwise	0.02	0.14	0	1
Single	1=single, 0=otherwise	0.14	0.34	0	1
Divorced	1=divorced, 0=otherwise	0.08	0.27	0	1
Widowed	1=widowed, 0=otherwise	0.07	0.26	0	1
Urban region	1=living in urban region, 0=otherwise	0.48	0.50	0	1
Undergoing urbanization	1=living in region undergoing urbanization, 0=otherwise	0.30	0.46	0	1
Rural area	1=living in rural area, 0=otherwise	0.22	0.42	0	1
Socioeconomic status					
Basic track	1=secondary general school leaving certificate or no degree, 0=otherwise	0.33	0.47	0	1
Intermediate track	1=intermediate school degree, 0=otherwise	0.28	0.45	0	1

Continued on next page...

... table 2.A.1 continued

Variable	Description	Mean	SD	Min.	Max
Academic track	1=leaving certificate from vocational school or college entry exam, 0=otherwise	0.35	0.48	0	1
Vocational training	1=vocational degree, 0=otherwise	0.72	0.45	0	1
University degree	1=university degree, 0=otherwise	0.31	0.46	0	1
Household income	Log. net equivalent hh income	10.02	0.59	0	12.84
Employed	1=full-time or part-time work, 0=otherwise	0.58	0.49	0	1
Not employed	1=not working, 0=otherwise	0.38	0.49	0	1
Unemployed	1=registered unemployed, o=otherwise	0.04	0.20	0	1
Social activities					
Volunteer work	1=volunteering at least monthly, 0=less often or never	0.22	0.41	0	1
Political participation	1=participating in political activities at least monthly, 0=less often or never	0.04	0.19	0	1
Religious participation	1=attending chuch or religious events at least monthly, 0=less often or never	0.19	0.39	0	1
Visit neighbors/friends	1=visiting neighbors and friends at least monthly, 0=less often or never	0.75	0.43	0	1
Visit family	1=visiting relatives at least monthly, 0=less often or never	0.76	0.43	0	1
Leisure activities					
Entertainment attendance	1=visiting cinemas, pop concerts, discos at least monthly, 0=less often or never	0.15	0.36	0	1

Continued on next page...

... table 2.A.1 continued

Variable	Description	Mean	SD	Min.	Max
Sports participation	1=excercising weekly, 0=less often	0.44	0.50	0	1
Go eating/drinking	1=going out for a meal or drink at least monthly, 0=less often or never	0.55	0.50	0	1
Excursions	1=going on excursions or trips at least monthly, 0=less often or never	0.29	0.45	0	1
TV consumption	1=watching TV or video daily, 0=less often	0.83	0.37	0	1
Computer use	1=using computer weekly, 0=less often	0.57	0.49	0	1
Artistic activities	1=pursuing artistic activities at least monthly, 0=less often or never	0.24	0.42	0	1
Garden work	1=doing garden work, hand crafts or repairing at least monthly, 0=less often or never	0.67	0.47	0	1
Car repair	1=doing car repair or maintenance at least monthly, 0=less often or never	0.25	0.43	0	1
Sport event attendance	1=attending sport events at least monthly, 0=less often or never	0.11	0.31	0	1
Health behavior					
Body-mass index (BMI)	BMI=weight in kgs/height in ms	26.32	4.62	16.10	67.20
Smoking status	1=currently smokes, 0=otherwise	0.23	0.42	0	1
Alcohol consumption	1=drinking alcohol regularly, 0=drinking alcohol less often	0.22	0.42	0	1
Nutrition	1=keeping healthy diet, 0=otherwise	0.43	0.50	0	1
Personality traits					
Openness	Score on the openness scale	4.55	1.07	1	7

Continued on next page...

... table 2.A.1 continued

Variable	Description	Mean	SD	Min.	Max
Conscientiousness	Score on conscientiousness scale	5.95	0.79	2	7
Extraversion	Score on the extraversion scale	4.83	1.01	1.33	7
Agreeableness	Score on the agreeableness scale	5.43	0.86	1.83	7
Neuroticism	Score on the neurticism scale	3.77	1.12	1	7
Early exposure to the arts					
Musical activity in youth	1=plyaed a musical instrument during youth, 0=otherwise	0.33	0.47	0	1
Sports participation in youth	1=did sports during youth, 0=otherwise	0.55	0.50	0	1
Parents: basic track	1=secondary general schooling leaving certificate or no degree, 0=otherwise	0.73	0.44	0	1
Parents: intermediate track	1=intermediate school degree, 0=otherwise	0.17	0.38	0	1
Parents: academic track	1=leaving certificate from vocational school or college entry exam, 0=otherwise	0.08	0.27	0	1
Parents: vocational training	1=vocational degree, 0=otherwise	0.55	0.50	0	1
Parents: university degree	1=university degree, 0=otherwise	0.06	0.23	0	1
Large city	1=lived in large city, 0=otherwise	0.25	0.43	0	1
Medium city	1=lived in medium city, 0=otherwise	0.18	0.39	0	1
Small city	1=lived in small city, 0=otherwise	0.21	0.41	0	1
Countryside	1=lived in the countryside, 0=otherwise	0.35	0.48	0	1
Number of siblings	Number of siblings	1.91	1.69	0	13

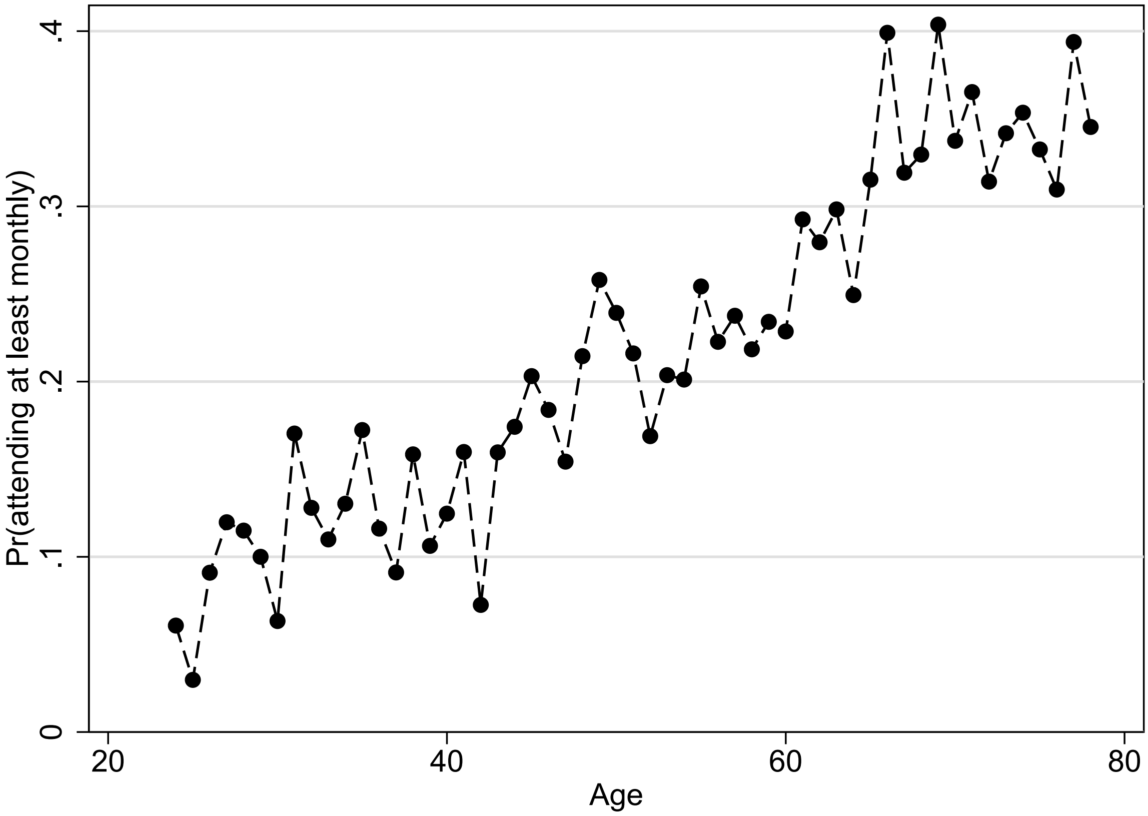
Note: Number of individuals=4,158.

Source: Own calculations based on SOEP v29.

Table 2.A.2: SOEP SF-12 health scales

Physical health scale (*pcs*)	
Physical functioning (2)	*Scale: 1 (greatly) to 3 (not at all)*
State of health affects ascending stairs	When you ascend stairs, i.e. go up several floors on foot: Does your state of health affect you greatly, slightly or not at all?
State of health affects tiring tasks	And what about having to cope with other tiring everyday tasks, i.e. when one has to lift something heavy or when one requires agility: Does your state of health affect you greatly, slightly or not at all?
Role physical (2)	*Scale: 1 (always) to 5 (never)* Please think about the last four weeks. How often did it occur within this period of time, that due to physical health problems ...
Achieved less due to health last 4 weeks	you achieved less than you wanted to at work or in everyday tasks?
Limited due to health last 4 weeks	you were limited in some form at work or in everyday tasks?
Bodily pain (1)	*Scale: 1 (always) to 5 (never)* Please think about the last four weeks. How often did it occur within this period of time, that due to physical health problems thad you had strong physical pains?
General health (1)	*Scale: 1 (very good) to 5 (bad)* How would you describe your current health?
Mental health scale (*mcs*)	
Vitality (1)	*Scale: 1 (always) to 5 (never)* During the last four weeks, how often did you: feel energetic?
Social functioning (1)	*Scale: 1 (always) to 5 (never)* During the last four weeks, how often did you: feel that due to physical and mental health problems your were limited socially, that is, in contact with friends, acquaintances, or relatives?
Role emotional (2)	*Scale: 1 (always) to 5 (never)* During the last for weeks, how often did you: feel that due to mental health or emotional problems...
Achieved less due to mental health the last 4 weeks	you achieved less than you wanted to at work or in everyday activities?
Less thorough due to health last 4 weeks	you carried out your work or everyday tasks less thoroughly than usual?
Mental health (2)	*Scale: 1 (always) to 5 (never)* During the last four weeks, how often did you:...
Run-down, melancholy last 4 weeks	feel down and gloomy?
Well-balanced last 4 weeks	feel calm and relaxed

Source: SOEP v29.



Source: Own calculations based on SOEP v29.

Figure 2.A.1: Predicted probability of cultural-event attendance (at least monthly) by age.

Table 2.A.3: The correlates of cultural participation (propensity score model)

	Coef.	SE
Female	0.006	(0.015)
Log(Household size)	0.010	(0.022)
Log(Number of children)	-0.023	(0.023)
Separated	0.018	(0.047)
Single	0.049*	(0.026)
Divorced	0.052**	(0.025)
Widowed	0.003	(0.030)
Urban region	0.010	(0.016)
Undergoing urbanization	-0.022	(0.018)
Intermediate track	0.045***	(0.017)
Academic track	0.059***	(0.021)
Vocational training	-0.053***	(0.016)
University degree	0.021	(0.018)
Log(Household income)	0.059***	(0.013)
Not employed	0.011	(0.019)
Unemployed	0.107***	(0.035)
Volunteer work	0.038***	(0.015)
Political participation	0.156***	(0.027)
Religious participation	0.056***	(0.016)
Visit neighbors/friends	0.061***	(0.016)
Visit family	0.005	(0.015)
Entertainment attendance	0.145***	(0.016)
Sports participation	0.038***	(0.013)
Go eating/drinking	0.082***	(0.014)
Excursions	0.061***	(0.013)
TV consumption	-0.012	(0.016)
Computer use	0.013	(0.015)
Artistic activities	0.097***	(0.014)
Garden work	-0.011	(0.014)
Car repair	-0.022	(0.016)
Sport event attendance	0.018	(0.019)
Body-mass index (BMI)	-0.004**	(0.002)
Smoking status	-0.046***	(0.016)
Alcohol consumption	-0.001	(0.015)
Nutrition	-0.030**	(0.013)
Openness (std.)	0.034***	(0.007)
Conscientiousness (std.)	-0.011*	(0.006)

Continued on next page...

... table 2.A.3 continued

	Coef.	SE
Extraversion (std.)	-0.006	(0.007)
Agreeableness (std.)	0.002	(0.007)
Neuroticism (std.)	-0.013**	(0.007)
Musical activitiy in youth	0.008	(0.013)
Sports participation in youth	0.006	(0.013)
Parents: basic track	0.010	(0.017)
Parents: academic track	0.041	(0.027)
Parents: vocational training	-0.001	(0.014)
Parents: university degree	-0.019	(0.033)
Large city	0.015	(0.018)
Medium city	-0.001	(0.017)
Small city	-0.007	(0.017)
Number of siblings	-0.002	(0.004)

Note: Number of individuals=4,158. Standard errors in parentheses. std. = z-standardized. The estimates are based on a probit model with the binary cultural participation indicator as the dependent variable. The coefficients measure marginal effects evaluated at the mean of the control variables. For reasons of clarity and comprehensibility, the coefficients on the age dummies are excluded and illustrated in Figure (2.A.1). Significance levels: *** $p<0.01$, ** $p<0.05$, * $p<0.1$.

Source: Own calculations based on SOEP v29.

Table 2.A.4: Covariate balance statistics

	Unmatched			Matched		
Variable	Treated	Control	%Bias	Treated	Control	%Bias
Female	0.52	0.51	1.0	0.52	0.53	-2.5
Household size	2.36	2.58	-18.7	2.37	2.41	-3.6
Number of children	0.27	0.45	-23.2	0.27	0.31	-4.7
Separated	0.02	0.02	0.2	0.02	0.02	-0.6
Single	0.13	0.14	-2.1	0.13	0.14	-2.5
Divorced	0.08	0.08	0.0	0.08	0.08	-1.2
Widowed	0.05	0.06	-4.4	0.05	0.06	-3.3
Urban region	0.56	0.46	21.8	0.56	0.58	-3.5
Undergoing urbanization	0.24	0.31	-15.1	0.25	0.24	2.1
Intermediatde track	0.24	0.30	-12.3	0.25	0.27	-6.5
Academic track	0.57	0.29	57.9	0.56	0.54	4.2
Vocational training	0.58	0.77	-42.4	0.59	0.60	-3.1
University degree	0.55	0.25	64.9	0.54	0.52	3.5
Household income	59016.75	42549.37	41.3	58813.28	59981.77	-2.9
Not employed	0.43	0.35	16.6	0.42	0.40	5.9
Unemployed	0.03	0.05	-10.5	0.03	0.03	1.3
Volunteer work	0.34	0.19	33.4	0.33	0.35	-4.0
Political participation	0.10	0.02	32.1	0.08	0.09	-3.8
Religious participation	0.26	0.17	21.9	0.25	0.25	-0.3
Visit neighbors/friends	0.86	0.73	32.8	0.85	0.85	1.7
Visit family	0.76	0.76	-0.4	0.76	0.76	-1.0
Entertainment attendance	0.31	0.11	51.7	0.30	0.30	0.5
Sports participation	0.64	0.39	51.1	0.64	0.63	0.4
Go eating/drinking	0.78	0.49	62.6	0.78	0.77	1.3
Excursions	0.49	0.23	56.6	0.48	0.48	0.3
TV consumption	0.78	0.85	-17.3	0.79	0.78	1.9
Computer use	0.69	0.56	28.8	0.69	0.70	-2.7
Artistic activities	0.45	0.18	60.7	0.43	0.44	-1.2
Garden work	0.70	0.67	5.3	0.70	0.68	3.4
Car repair	0.20	0.27	-16.1	0.20	0.22	-2.7
Sport event attendance	0.14	0.11	8.7	0.13	0.14	-1.3
Body-mass index (BMI)	25.53	26.59	-23.8	25.55	25.45	2.3
Smoking status	0.15	0.26	-28.0	0.15	0.16	-3.5
Alcohol consumption	0.28	0.21	18.3	0.28	0.28	0.1
Nutrition	0.30	0.47	-34.7	0.31	0.30	0.6

Continued on next page...

... table 2.A.4 continued

	Unmatched			Matched		
Variable	Treated	Control	%Bias	Treated	Control	%Bias
Openness	5.00	4.42	57.2	4.98	5.02	-4.1
Conscientiousness	5.92	5.97	-6.6	5.91	5.92	-1.3
Extraversion	4.95	4.79	16.2	4.94	4.95	-0.8
Agreeableness	5.48	5.41	8.4	5.47	5.48	-0.8
Neuroticism	3.59	3.82	-20.6	3.60	3.63	-2.4
Musical activity in youth	0.45	0.28	34.3	0.44	0.46	-4.0
Sports participation in youth	0.62	0.54	14.4	0.61	0.60	3.2
Parents: intermediate track	0.24	0.16	19.0	0.23	0.23	1.1
Parents:academic track	0.14	0.06	25.0	0.13	0.14	-4.3
Parents: vocational training	0.57	0.55	4.5	0.57	0.58	-1.2
Parents: university degree	0.08	0.05	14.0	0.08	0.10	-9.0
Medium city	0.20	0.18	5.0	0.19	0.19	1.0
Small city	0.20	0.22	-3.7	0.21	0.21	-0.6
Countryside	0.29	0.37	-18.0	0.29	0.28	1.5
Number of siblings	1.73	1.95	-12.9	1.73	1.76	-1.7

Note: Number of individuals=4,158. All continuous variables are measured on their original scale.
Source: Own calculations based on SOEP v29.

Derivation of the Sensitivity Analysis Parameters

The parameters $p_{ij} = Pr(U = 1|T = i, Y = j)$ corresponding to given values of d and s can be derived by solving a system of equations. To make the analysis more tractable, the values for $Pr(U = 1)$ and $p_{11} - p_{10}$ can be fixed beforehand. However, it is assumed that the influence of these items on the selection process and outcomes is negligible (Nannicini et al., 2007). We therefore define $Pr(U = 1) = 0.5$ and $p_{11} - p_{10} = 0$. As an example, assume that $d = 0.1$ and $s = 0.1$. The system of equations then reads as follows:

$$0 = p_{11} - p_{10}$$

$$\begin{aligned} Pr(U=1) = 0.5 &= p_{11} \cdot Pr(Y=1|T=1) \cdot Pr(T=1) + p_{10} \cdot Pr(Y=0|T=1) \cdot Pr(T=1) \\ &+ p_{01} \cdot Pr(Y=1|T=0) \cdot Pr(T=0) + p_{00} \cdot Pr(Y=0|T=0) \cdot Pr(T=0) \end{aligned}$$

$$d = 0.1 = p_{01} - p_{00}$$

$$\begin{aligned} s = 0.1 &= p_{1.} - p_{0.} \\ &= p_{11} \cdot Pr(Y=1|T=1) + p_{10} \cdot Pr(Y=0|T=1) \\ &- p_{01} \cdot Pr(Y=1|T=0) - p_{00} \cdot Pr(Y=0|T=0) \end{aligned}$$

Table 2.A.5: Sensitivity analysis: unobserved confounder similar to observed covariates (mental health only)

	Out. effect		Sel. effect			
	d	Γ	s	Λ	Coef.	SE
Sex	-0.10	0.7	0.00	1.1	0.639	(0.361)
Age	0.10	1.6	0.16	1.9	0.433	(0.368)
Household size	-0.06	0.8	-0.12	0.6	0.529	(0.366)
Number of children	-0.07	0.7	-0.11	0.5	0.506	(0.368)
Married	0.05	1.3	0.01	1.1	0.624	(0.361)
Separated	0.00	0.8	0.00	1.3	0.634	(0.361)
Single	-0.03	0.8	0.00	1.0	0.633	(0.361)
Divorced	-0.01	1.0	0.00	1.1	0.633	(0.361)
Widowed	-0.01	0.8	-0.02	0.7	0.621	(0.362)
Urban region	0.03	1.2	0.11	1.6	0.565	(0.365)
Undergoing urbanization	0.00	1.0	-0.07	0.7	0.615	(0.364)
Rural area	-0.03	0.8	-0.04	0.8	0.606	(0.361)
Basic track	0.01	1.0	-0.22	0.3	0.612	(0.382)
Intermediate track	-0.03	0.9	-0.05	0.8	0.607	(0.363)
Academic track	0.02	1.1	0.29	3.6	0.544	(0.394)
Vocational training	-0.02	0.9	-0.20	0.4	0.559	(0.384)
University degree	0.05	1.3	0.31	4.0	0.381	(0.402)
Household income	0.09	1.5	0.29	3.4	0.334	(0.390)
Employed	-0.02	0.9	-0.05	0.8	0.614	(0.361)
Not employed	0.03	1.1	0.07	1.3	0.595	(0.362)
Unemployed	-0.01	0.8	-0.02	0.6	0.620	(0.362)
Volunteer work	0.01	1.1	0.15	2.2	0.570	(0.372)
Political participation	0.00	0.8	0.08	5.4	0.647	(0.375)
Religious participation	0.00	1.0	0.09	1.7	0.590	(0.365)
Visit neighbors/friends	0.05	1.3	0.14	2.4	0.531	(0.371)
Visit family	0.01	1.0	0.00	1.0	0.631	(0.361)
Entertainment attendance	-0.01	1.0	0.20	4.0	0.613	(0.391)
Sports participation	0.04	1.2	0.24	2.8	0.495	(0.384)
Go eating/drinking	0.07	1.3	0.30	4.0	0.421	(0.395)
Excursions	0.07	1.5	0.27	3.5	0.348	(0.390)
TV consumption	0.03	1.3	-0.07	0.6	0.639	(0.364)
Computer use	0.01	1.1	0.14	1.9	0.578	(0.369)
Artistic activities	0.00	1.0	0.27	4.0	0.589	(0.401)
Garden work	0.00	1.0	0.03	1.2	0.624	(0.362)
Car repair	0.01	1.1	-0.07	0.7	0.627	(0.364)

Continued on next page...

... table 2.A.5 continued

	Out. effect		Sel. effect			
	d	Γ	s	Λ	Coef.	SE
Sport-event attendance	0.03	1.3	0.03	1.4	0.604	(0.362)
Body-mass index (BMI)	0.01	1.1	-0.09	0.7	0.618	(0.364)
Smoking status	-0.03	0.9	-0.11	0.5	0.569	(0.368)
Alcohol consumption	0.06	1.5	0.08	1.5	0.551	(0.364)
Nutrition	-0.04	0.9	-0.17	0.5	0.524	(0.374)
Conscientiousness	0.12	1.7	-0.03	0.8	0.677	(0.362)
Openness	0.04	1.2	0.22	2.6	0.520	(0.384)
Extraversion	0.07	1.3	0.06	1.3	0.584	(0.363)
Agreeableness	0.11	1.6	0.03	1.1	0.613	(0.362)
Neuroticism	-0.27	0.3	-0.09	0.8	0.456	(0.362)
Musical activitiy in youth	-0.02	0.9	0.17	2.2	0.649	(0.378)
Sports participation in youth	0.03	1.1	0.08	1.4	0.592	(0.363)
Parents basic track	0.01	1.1	-0.15	0.5	0.597	(0.372)
Parents intermediate track	0.00	1.0	0.08	1.8	0.602	(0.365)
Parents academic track	-0.01	0.9	0.07	2.6	0.633	(0.367)
Parents vocational training	-0.01	1.0	0.03	1.1	0.629	(0.361)
Parents university degree	0.00	1.0	0.04	1.9	0.628	(0.364)
Large city	0.01	1.1	0.08	1.5	0.591	(0.363)
Medium city	0.01	1.0	0.02	1.2	0.625	(0.361)
Small city	0.00	1.0	-0.02	0.9	0.629	(0.361)
Countryside	-0.01	0.9	-0.08	0.7	0.600	(0.363)
Number of siblings	0.00	1.0	-0.05	0.8	0.626	(0.362)

Note: Number of individuals=4,158. Robust standard errors from weighted linear regressions in parentheses. d measures the difference in the incidence of the unobserved variable U between healthy and unhealthy individuals in the absence of cultural participation ($p_{01} - p_{00}$). s measures the difference in the incidence of U between participants and non-participants ($p_{1.} - p_{0.}$). Γ and Λ are odds ratios obtained from a logit regression including U and all covariates as independent variables, and quantify the effects of U on the untreated outcome and participation, respectively.

Source: Own calculations based on SOEP v29.

CHAPTER 3

Social Capital, Caregiving and Mental Health

3.1 Introduction

For several years, great effort has been devoted to the study of the health effects of informal care – that is care provided to a loved one at home — and the implications for health care and labor markets (for a recent review, see Bauer and Sousa-Poza, 2015). With respect to psychological health outcomes, empirical research suggests that caregivers more frequently report greater degrees of depression, anxiety, or stress and exhibit lower levels of subjective well-being than non-caregivers (e.g. Schulz and Sherwood, 2008). Recent evidence from German individual-level data even indicates a causal impact of informal care duties on mental health (Schmitz and Westphal, 2015). The repercussions of poor caregiver health on health care costs and labor markets can thus be substantial. This is why many developed countries have introduced and expanded public support programs to address the well-being of informal carers (for an international overview of caregiver support policies, see Colombo et al. (2011, ch. 3)). For instance, German health insurance and long-term care (LTC) providers offer respite care, training and counseling, and coordinated information services for family carers. Although these services are widely known, and informal caregivers have a legal right to obtain them, recent surveys among the insured suggest that utilization rates are rather low (Robert Koch Institute, 2015).

This chapter looks at an alternative source of caregiver support, that is the individual's social capital. Specifically, we ask whether caregivers with stronger social ties have better psychological health than carers with fewer social connections. The basic hypothesis states that social capital buffers the negative mental health effects of caregiving. This is because social bonds such as family, friends, or neighbors may offer resources for emotional, instrumental, and informational support (e.g. Cohen and Wills, 1985; Thoits, 2011), which can reduce the psychological burden of caregiving.[9]

This research is related to studies documenting how individuals from different subgroups of the population react to caregiving responsibilities. A great deal of this research has examined gender differences, even though female and male caregivers display similar mental health states when caregiving stressors (e.g. care receiver's health, or hours of care) and social resources are accounted for (Pinquart and Sörensen, 2006). Some studies have investigated the interaction between informal caring and employment with respect to psychological health. The protective effects of working may result from the social resources involved in gainful employment, e.g. support received from colleagues or employers (Hansen and Slagsvold, 2015). However, informal caring along with full-time work may also imply a double burden and exacerbate the negative health consequences of caregiving (Schmitz and Stroka, 2013). This study also adds to prior research that has addressed heterogeneous effects of informal care depending on marital status or whether the caregiver cohabits with the care recipient (e.g. Hansen et al., 2013; Van den Berg and Ferrer-i Carbonell, 2007).

Another strand of (economic) literature examines the insurance effect of social participation with respect to adverse life events such as divorce, unemployment and negative income shocks

[9]Note that we use the terms social capital, social ties, social bonds, social networks, social interaction, social participation etc. interchangeably, acknowledging the different theoretical connotations in the literature.

(Clark and Lelkes, 2006; Dehejia et al., 2007; Winkelmann, 2009). The general argument is that social capital may reduce the negative well-being consequences of these economic shocks. As a result, individuals with more social capital, who have higher utility in the case of adversity than those with less social capital, may rely less on governmental or social security benefits.

Generally, we contribute to the caregiving literature in the following ways: First, this is, to the best of our knowledge, the first study which uses rich survey data from the larger general population to analyze the moderating role of social capital with respect to caregiver mental well-being. A key limitation of previous research is that it is based on small samples, with a focus on caregivers and health care professionals (see, for example, the reviews and analyses by Cannuscio et al., 2004; Rodakowski et al., 2012; Barrow and Harrison, 2005). Second, we provide a more in-depth analysis of the interaction between caregiving and social capital. We examine whether and to what extent care intensity (i.e. hours of care) and different types of social interactions influence the buffering process. Third, this study contains a variety of sensitivity checks to assess how observed characteristics, correlated with social capital, and the caregiving context explain the moderating role of strong social ties. Furthermore, we also estimate buffering effects using a matched sample of carers and non-carers where each person theoretically has the same probability of caregiving. Fourth, the empirical models include an extensive list of personal and household characteristics that might simultaneously influence the decision to provide informal care, social activities, and mental health. In addition, we take unobserved individual effects into account that do not change over time, but which can influence the estimated relationships.

We use individual-level data from the German Socio-Economic Panel (SOEP) study. Our baseline estimates suggest that social capital (measured basically as the equally weighted sum of formal and informal social activities) weakens the negative association between informal care and mental health. A one-unit increase in the social capital index, which is measured roughly on the same scale as caring status, reduces the negative correlation between caregiving and mental health by 54%. Looking at the mental health subscales, we find that social capital improves vitality, and alleviates depressive symptoms and perceived time pressure among caregivers. Furthermore, caregivers with high time commitments or who participate regularly in voluntary organizations experience significant buffering effects.

Sensitivity analyses suggest that the buffering role of social capital cannot be explained by the moderating influence of observable factors correlated with social capital. The findings are also robust to the inclusion of caregiving context variables (care-receiver health, other sources of support, and the relationship between the caregiver and the dependent person) as potential buffering mechanisms. However, the results might be driven by a comparison of individuals that differ widely in terms of observable characteristics and the propensity to become a caregiver, respectively. This is a problem in linear regression models which account for observable characteristics but include control observations with equal weight. Thus, non-caregivers with a close to zero probability of caring may contribute considerably to the estimates. Indeed, in a fixed effects regression using a matched sample of caregivers and non-caregivers, who have roughly the same probability of caring, the coefficient for the moderator effect decreases substantially.

The moderating role of social capital could explain the low utilization rates regarding caregiver support services. Persons with higher social capital tap into their social network to obtain assistance rather than relying on public support programs. Many individuals may simply have a preference for informal help because the care receiver rejects care provided by strangers, the perceived cost of support services are too high, or they do not know how and where to obtain these services (Jacobs et al., 2016). Policies to promote caregiver well-being should therefore also involve measures to strengthen social interaction and participation in the neighborhood or the community. As argued by Putnam (2001), engagement in social activities is one of the main drivers of social capital and trust and contributes to the formation of social networks that provide and foster norms of mutual assistance and reciprocity. Thus, social participation provides the fundamental resources to which individuals can turn to under demanding situations, such as providing informal care to a family member.

The remainder of this chapter is organized as follows: Section (3.2) reviews the theoretical and empirical literature regarding the buffering role of social capital in the context of informal care. Section (3.3) describes the data and outlines the econometric method used for our analysis. In Section (3.4), we present and discuss the main estimation results. Section (3.5) provides several sensitivity checks. Section (3.6) summarizes the main findings and concludes.

3.2 Related Literature

This study largely builds on sociological and psychological research regarding the stress-buffering role of social capital, which have also found their way into the economic literature. Caregiving is generally understood as a stress process, which involves a variety of characteristics of the caregiver, the care recipient and family background that shape the experience of informal care provision. An important contextual factor is a person's social network or support received from others, because it may alter the way how the caregiving burden translates into mental and physical health problems (Pearlin et al., 1990). In their seminal paper, Cohen and Wills (1985) make this assumption more explicit and advance the *stress-buffering hypothesis*. After reviewing evidence on the relationship between stress, social support and well-being, they argue that integration in social networks can reduce the negative health consequences of taxing circumstances. Specifically, support received from others (family, friends, neighbors etc.) can improve health by altering the assessment of the stressful event or situation at several occasions.

The extant research proposes a variety of mechanisms that explain why and how social ties can reduce stress and mental health problems. The comprehensive reviews by Kawachi and Berkman (2001) and Thoits (2011) suggest that the buffering effect mainly operates through support received from social contacts. They basically distinguish between three types of social support: emotional, instrumental, and informational support. Emotional support involves social-psychological mechanisms that can improve psychological well-being in demanding circumstances. By expressing understanding, concern and care for the caregiver's problems, social ties can reduce the burden and stress associated with providing informal care. Spending time with others in social or leisure activities may also serve as an opportunity for the caregiver to distract oneself from the potentially difficult caregiving duty. Instrumental support comprises practical assistance or financial aid. Thus, caregivers may receive financial payments or benefits in-kind from their relatives, friends, or neighbors; or social contacts may simply assume caregiving tasks and responsibilities, for example when the primary caregiver is working. These factors may improve psychological well-being by easing the caregiver's tasks and facilitating the compatibility of informal care with other activities such as work or leisure.

Social ties may also act a valuable resource for information and advice. According to Durlauf and Fafchamps (2005), the individual's social capital is an important resource for information. The health economic literature frequently highlights the role of social capital with respect to information problems on health care markets. Social interactions may improve the knowledge about the health system's institutional details and the availability and suitability of medical and psychological treatments (e.g. Deri, 2005; Folland, 2008).

Generally, previous studies find a positive association of social capital with caregivers' mental well-being (for reviews, see for example Cannuscio et al., 2004; Rodakowski et al., 2012). A major shortcoming of these studies is that the majority is based on small-scale samples paying attention to specific subgroups of the population (e.g. caregivers only or nurses), and that they also omit many characteristics that might confound the relationship between social contacts and health among caregivers. Only a handful of studies exist that use population-level data and investigate the buffering effect employing a regression framework. Using a cross-section of the Nurses' Health Study (NHS) from the U.S., Cannuscio et al. (2004) have reported that a higher caring workload is associated with an increased risk of depressive symptoms. This relationship was more pronounced among women with fewer social ties. Barrow and Harrison (2005) have analyzed the potential role of neighborhood attachment as a modifier of the caregiving-health nexus. They have found that caregivers with a higher sense of belonging to the community experience less physical and mental health problems than caregivers who felt more alienated. Similarly, Carpiano (2008) has examined how actual neighborhood attachment influences the buffering effects of social capital. They have found that stronger ties to the neighborhood promote the buffering of the caregiver burden. This provides evidence that strong social connections facilitate access to the tangible and intangible resources available in the community, which could improve caregiver health and well-being.

Empirical economic research regarding the buffering hypothesis is rather scarce and, to the best of our knowledge, absent in the context of caregiving. Nevertheless, the stress-buffering property of social capital might also be economically relevant. From an economic perspective, individuals invest in their social capital to acquire market and non-market based resources (Glaeser et al., 2002). By building up social capital through social participation, for instance through civic or voluntary engagement, individuals obtain valuable resources which they can harness when they need them (see also Coleman, 1988). Probably the first econometric test of stress buffering was conducted by Clark and Lelkes (2006), who have analyzed whether religious involvement mitigates the negative effect of adverse life events on subjective well-being. They have found significant moderating effects of religiosity for unemployment and marital dissolution. Dehejia et al. (2007) extend the analysis to income and consumption shocks, respectively.

They have shown that regular religious participation mitigates the negative happiness effect of adverse income shocks. Winkelmann (2009) has also tested the buffering hypothesis among German individuals who became unemployed. However, he finds no significant buffering effect of social capital (measured by a variety of formal and informal social activities) on the unemployed's subjective well-being. Generally, individuals with more social capital and with greater happiness or utility levels in the event of economic shocks may have lower demand for governmental or social security benefits, since they can rely on the resources embedded in their social network.

We argue that social capital can (partially) insure the individual against the negative psychological consequences of informal care provision. It does so by providing resources for emotional and practical support. It may also involve valuable information about resources that directly or indirectly address both the care receiver's and the caregiver's health and well-being.

3.3 Data and Empirical Approach

To examine whether and how the negative association between caregiving and mental health is modified by social capital, we use data on individuals and families from the German Socio-Economic Panel (SOEP) study (Wagner et al., 2007).[10] The SOEP is a longitudinal survey of German households running since 1984 providing information on employment, income, subjective well-being and other characteristics. Recent survey years have included more detailed measures of self-rated health and time use, which also involve informal care provision and social activities.

3.3.1 Measurement of Variables

Mental Health

As the main outcome, we use the mental component summary scale (*MCS*) extracted from the SOEP version of the 12-item short-form health survey (SF-12), which has been included biennially since 2002 (see also Andersen et al., 2007). The SF-12 questionnaire contains a variety of questions that cover the respondents health-related quality of life. Twelve items are aggregated into eight subscales and the two major dimensions "mental health" and "physical health" by means of factor analysis. The mental-health scale of the SF-12 survey in the SOEP sums the subcomponents *vitality* (*VT*), *social functioning* (*SF*), *role emotional* (*RE*), and *mental health* or *depressive symptoms* (*MH*).[11] The corresponding items refer to the individual's self-reported health status in the last four weeks, and are assessed on a five-point scale (from 1=always to 5=never). *Vitality* measures how often the respondent had felt energetic. *Social functioning* reflects to what extent mental health problems limited social contacts, whereas *role emotional* provides information on whether and how strongly mental health problems interfered with work

[10]Socio-Economic Panel (SOEP), data for years 1984-2013, version 30, SOEP, 2015, doi:10.5684/soep.v30.

[11]See Table (3.A.1) in the appendix for further details on the mental component summary scale.

and daily activities. Finally, *mental health* reflects the frequency of depressive symptoms, notably sadness or agitation.

What is more, we also consider perceived time pressure or stress (TS) as a dependent variable, which has rarely been done in previous studies. We argue that it is an important outcome in the context of caregiving, since providing informal care interferes with other activities such as leisure and work. The question on perceived time pressure uses the same categories as the other scales, and respondents are asked to assess how often they had felt pressed for time in the last four weeks. We expect that caregivers report being more frequently pressed for time than non-caregivers, but that caregivers with more social capital feel less time-pressed than their counterparts with fewer social bonds.

We will use the MCS, its four subscales and the time stress variable as outcomes in the regression analyses. The dependent variables are standardized so that they have a mean of zero and a standard deviation equal to one. Higher scores reflect improved health status.

Caregiving Status

Caregiving status is computed based on information regarding caregiving relationship, hours and cohabitation status. We create a dummy variable *Caregiver* which is equal to one when the respondent reports at least one hour of care per week and is the main caregiver to a care-dependent person who lives in the same household. Confining the sample to cohabiting caregivers allows us to include the caregiving context characteristics into the analysis. This becomes particularly important when we examine alternative explanations for the moderating role of social capital. For example, care-recipient health — which may also be correlated with the caregiver's social capital — may also influence the perceived burden of informal care.

With the data at hand, we can also explore the importance of different caregiving intensities. We create a categorical variables which reflects increasing hours of care per week and which is loosely based on the eligibility criteria of the German LTC insurance (for a general overview of the German LTC system, see Zuchandke et al., 2011). Basically, persons who claim benefits from the LTC insurance fund need to undergo a medical examination which assesses the care needs of the recipient. After evaulation is completed the care recipient is usually assigned to one of three care levels that reflect increasing need and help in terms of (instrumental) activities of daily living. It also involves an assessment of time necessary for nursing care which serves as a basis for our caregiving intensity or hours variable. The categories are defined as follows: 0 (non-caregivers), 1-9, 10-20 , 21-34, and 35 and more hours per week. We expect that there is a negative correlation between weekly informal care workload and mental health. Social contacts might be particularly useful for caregivers with a high time commitment (Cannuscio et al., 2004).

Social Capital

To define individual-level social capital, we follow Putnam's (2001) approach. As a proxy for the stock of social capital, we use the individual's formal and informal social activities. On the one hand, individuals build up and maintain social capital by meeting with friends, going to the movies, or having a night out. On the other hand, civic engagement or participation in social organizations — such as political engagement, religious involvement, or volunteer work — may also constitute an important part of the individual's social capital (see also Glaeser et al., 2002). Similar concepts have recently been adopted by Bauernschuster et al. (2014) and Winkelmann (2009) who also use the SOEP data.

The SOEP study includes several formal and informal social activities, and the following variables are used[12]: the frequency of meeting with friends, relatives, or neighbors (*Social gatherings*); helping out friends, relatives or neighbors (*Helping*); involvement in a citizens' group, political party, or local government (*Political participation*); attending church or religious events (*Religious participation*); volunteer work in clubs or social services (*Volunteer work*); doing sports (*Sports participation*); going to cultural events (*Cultural attendance*); going to the movies, pop music concerts, dancing, disco, or sports events (*Entertainment attendance*); and artistic or musical activities (*Artistic activities*). Each item is assessed by means of four categories which reflect the frequency of the activity: never, less often, monthly, and weekly.

We follow the method proposed by Kling et al. (2007) to create a comprehensive social capital measure by adding up the z scores of the individual social capital components. We thus obtain a social-capital-index (*SOCI*) variable that assigns equal weight to each of its components and which reflects increasing levels of social capital.[13] In further analyses, however, we include the individual social capital components separately. They enter the regression models as dichotomous variables which reflect at least monthly participation in each activity.

Control Variables

We incorporate a variety of personal and household characteristics that might confound the relationship between mental health, caregiving, and social capital. The regression models include control variables for gender, age (and age squared), marital status (married, separated, single, divorced, widowed), and migration background (yes/no). Socioeconomic factors involve schooling (log. of years of education) and income (log. of net equivalized household income). We also take the respondent's labor market involvement and job characteristics into account. Employment is assessed using the individual's labor force status (employed, unemployed, not working). Furthermore, we include working hours and net hourly wages to control for time restrictions and opportunity costs related to the decision to become a caregiver and to socialize. We include a dummy variable for missing values when information on working hours and wages are not available or inapplicable. In these cases, the working hours and wages are zero.

To control for health-related selection, we include the widely used measure of self-assessed health (SAH). It consists of five categories reflecting different states of general health: bad, poor, satisfactory, good and very good. Although the direction is not always clear, empirical evidence generally demonstrates a strong correlation between self-reported health status on the one hand and the decision to provide informal care and social participation on the other hand (e.g. Bauer

[12]See Table (3.A.2) in the appendix for further details on question wording and response scales.

[13]Cronbach's alpha is equal to 0.67.

and Sousa-Poza, 2015; Rocco and Fumagalli, 2014). Omitting the respondent's health would therefore bias the associations between mental health, caregiving and social capital.

Additionally, we add indicators for household size (log. of number of persons) and children (log. of number of children). We differentiate between children at ages 0-6 on the one hand and at ages 7-16 on the other hand. While household size in general might indicate the availability of other sources of support, the existence of (younger) children rather reflects an additional burden for informal carers (e.g. Rubin and White-Means, 2009). Furthermore, we include measures for home ownership (tenant vs. home owner) and urbanization level (urban vs. rural area), both of which have been shown to correlate with individual-level social capital (Glaeser et al., 2002). Finally, we include the monthly amount of care allowance received by the household. To account for macroeconomic conditions and time trends, we incorporate both survey year and state dummy variables.

Summary and Descriptive Statistics

The final analysis sample consists of $70,680$ person-year observations resulting from $28,939$ individuals observed over a maximum period of five years. Social capital and mental health are measured in different survey years. Information on mental well-being is gathered in even survey years, whereas social activities are assessed in odd years. What is more, most of the social capital variables are not available in 2003. The data on caregiving, social capital and covariates therefore refer to the years 2001, 2005, 2007, 2009 and 2011, and we use the mental health outcomes one year later.

Table (3.1) displays the mean values and standard deviations of the variables used in the empirical analysis. Apparently, there are marked differences between non-caregivers and caregivers in terms of observed characteristics. For example, a raw comparison reveals that caregivers have significantly poorer general and mental health, and lower social capital than non-caregivers. Caregivers are more likely to be female, older and married than those who do not provide informal care. Caregivers also report fewer children than non-caregivers. Furthermore, caregiving seems to be associated with worse socioeconomic status, that is less years of schooling, lower income and a higher probability of joblessness. Working time and wages are also lower among informal carers. Intuitively, households with a care-dependent family member receive a higher monthly care allowance from the LTC fund.

3.3.2 Empirical Strategy

To examine the buffering mechanism induced by social capital in the context of caregiving, we perform linear regression analyses using the following model (surpressing the constant):

$$MH_{it} = \alpha Caregiver_{it} + \beta SOCI_{it} + \gamma \left(Caregiver_{it} \times SOCI_{it}\right) + \delta X' + u_i + \varepsilon_{it},$$

where i denotes the individual and t indicates the survey year. The coefficient α reflects the association between $Caregiver$ and the mental health outcome MH, which is assumed to be

Table 3.1: Descriptive statistics of estimation sample

	Noncaregivers		Caregivers		Diff.
	Mean	S.D.	Mean	S.D.	
MCS	0.01	0.10	-0.41	1.09	-0.41
SOCI	-2.93	4.84	0.05	4.70	2.98
Female	0.52	0.50	0.68	0.47	0.16
Age	49.38	16.95	60.73	13.89	11.34
Married	0.62	0.48	0.84	0.37	0.22
Separated	0.02	0.13	0.01	0.10	-0.01
Single	0.21	0.41	0.06	0.24	-0.15
Divorced	0.08	0.26	0.05	0.21	-0.03
Widowed	0.07	0.25	0.04	0.20	-0.03
Household size	2.62	1.25	2.72	1.03	0.09
No. of children age 0-6 in hh	0.16	0.47	0.07	0.31	-0.10
No. of children age 7-16 in hh	0.32	0.67	0.22	0.60	-0.10
Migration background	0.18	0.38	0.19	0.39	0.01
Degree of urbanization	0.66	0.48	0.60	0.49	-0.06
Years of education	12.15	2.67	11.49	2.44	-0.66
Household income	21,572.17	16,310.38	18,685.68	8,567.59	-2,886.49
Employed	0.58	0.50	0.28	0.45	-0.30
Unemployed	0.05	0.22	0.07	0.25	0.02
Not working	0.37	0.48	0.66	0.48	0.28
Working hours	22.21	21.52	8.31	15.65	-13.91
Net hourly wage	5.59	7.22	2.32	5.22	-3.26
Homeowner	0.51	0.50	0.60	0.49	0.07
SAH bad	0.04	0.18	0.06	0.25	0.03
SAH poor	0.14	0.34	0.21	0.41	0.07
SAH satisfactory	0.33	0.47	0.44	0.50	0.11
SAH good	0.41	0.49	0.25	0.43	-0.16
SAH very good	0.09	0.28	0.03	0.18	-0.05
Monthly care allowance	4.01	48.41	213.53	276.23	209.53
Observations	69,593		1,087		

Note: All differences between caregivers and non-caregivers are significant except for migration background.
Source: Own calculations based on SOEP v30.

negative; β measures the relationship between the social capital index $SOCI$ and mental health, which is expected to be positive; the parameter γ evaluates the interaction between $Caregiver$ and $SOCI$. Theory and previous empirical evidence suggests that the sign of the coefficient γ is positive. This implies that the negative association between caregiving and mental health is attenuated by increasing levels of social capital. The control variables are captured by the vector X'. The parameter u_i captures time-invariant individual effects, whereas ε_{it} reflects unobserved shocks that vary over time and individuals.

The $Caregiver$ and the $SOCI$ variable are centered around the mean and rescaled by dividing by two standard deviations of the original variable. This procedure facilitates comparisons between the dummy and the continuous variable since they are roughly measured on the same scale (Gelman, 2008). Otherwise, we would compare a discrete change from non-caregiver to caregiver with a one-unit change in the social capital index, which could understate the importance of social capital relative to caring. Standardization of the dichotomous variable is necessary because the caregiver distribution is highly skewed, with only 1.5% of observations in the caregiver group and a standard deviation of about 0.12. Using the unstandardized binary variable in the regression would therefore overstate the importance of caregiving relative to the standardized social capital index, which has a standard deviation of 0.5.

This regression produces unbiased estimates as long as unobserved differences between caregivers and non-caregivers, and socially active and inactive persons, are negligible. Unobserved preferences for helping and socializing with others, or an innate ability to perform various productive activities simultaneously, may drive the buffering effect by social capital. A further thread to the validity of estimates arises from the endogeneity or simultaneity between social activities and caring decisions. For instance, due to time restrictions, caregivers may reduce their social activities. This implies that those who participate frequently in social activities are less likely to be caregivers. The average mental health declinedue to caring would thus be less severe for individuals with stronger social ties. Nevertheless, the smaller decline in psychological well-being is not attributable to the protecting role of social capital, but rather due to non-caregivers predominantly selecting into social participation. Consequently, we would overestimate the buffering effect of social capital.

Ideally, to avoid confounding due to omitted variables and simultaneity, caregiving status and social activities should be as good as randomly assigned. With observational data, we need external variation at least for individual-level social capital. Using an instrumental variable for social capital in a caregiver-only sample, for instance, we could estimate a buffering effect that is independent of unobservable variables and caregiver status. However, convincing and valid instrumental variables for social capital are not available in our data. In this study, we estimate fixed effects models instead, to get rid of at least some portion of unobserved heterogeneity. That is, we subtract for each individual and each variable the corresponding time-averaged mean from the contemporary value. This means that any time-fixed variable, also the individual effects u_i, are eliminated because of this procedure. In fact, we are able to remove any unobserved differences across individuals that may influence our results as long as they are stable over time.

This approach does not completely solve the problems due to unobserved shocks and the potential endogeneity between social capital and caring. However, it still provides a major improvement compared to previous studies, such as Cannuscio et al. (2004) or Carpiano (2008), because we take a larger set of covariates and omitted variables bias into account. We therefore likely obtain less biased estimates than past research.

3.4 Results

3.4.1 Main Results

Table (3.2) reports the estimation results from the fixed effects (that is within-individual) regressions of mental well-being on caregiving status, the social capital index and the control variables. Alongside the main terms we also include the product between the standardized caregiver and social capital variables to capture the alleviating role of social ties in the caregiving-health relationship. The inclusion of the interaction term changes the interpretations of the main terms slightly. The parameter belonging to the caregiving indicator reflects the comparison between caregivers and non-caregivers among individuals with an average level of social capital. The coefficient of the social capital index assesses approximately the relationship with mental health for non-carers.

Table 3.2: The moderating role of social capital

	(1) MCS	(2) VT	(3) SF	(4) RE	(5) MH	(6) TS
Caregiver	-0.066***	-0.016	-0.039***	-0.060***	-0.075***	-0.068***
	(0.013)	(0.014)	(0.014)	(0.014)	(0.013)	(0.013)
SOCI	0.065***	0.109***	0.033**	0.023	0.058***	-0.009
	(0.014)	(0.014)	(0.014)	(0.014)	(0.014)	(0.014)
Caregiver×*SOCI*	0.049**	0.058***	0.019	0.011	0.044**	0.063***
	(0.020)	(0.021)	(0.023)	(0.020)	(0.019)	(0.019)
Observations	70,680	70,680	70,680	70,680	70,680	70,680

Note: Robust standard errors, clustered at the individual level, in parentheses. Dependent variables: *MCS* = mental component summary scale, *VT* = vitality, *SF* = social functioning, *RE* = role emotional, *MH* = depressive symptoms, *TS* = time stress. The caregiver dummy variable and the social-capital index are centered and standardized and have zero mean and a standard deviation of 0.5. All regression models include controls for age (and age squared), marital status (married, separated, single, divorced, widowed), household size, number of children at ages 0-6 and 7-16, degree of urbanization, migration background, schooling (years of education), income (net household income), employment status (employed, not working, unemployed), working hours, hourly net wage, ownership status (owner vs. tenant), states, and survey years. Individual fixed effects are also included. Significance levels: *** p<0.01, ** p<0.05, * p<0.1.

Source: Own calculations based on SOEP v30.

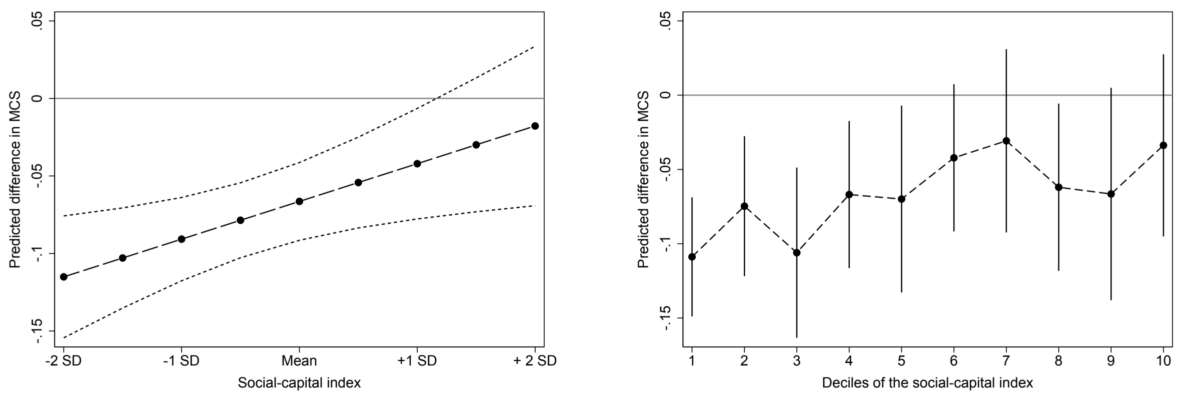
The estimates on both the main and the interaction variables generally go in the hypothesized direction. Column 1 shows the results for the mental health summary scale MCS, our primary health outcome. The parameter estimate for the caregiver dummy variable is -0.066 and significant, reflecting a negative relationship between informal care and mental health for individuals who display average levels of social capital.[14] The parameter for the social capital index is 0.065 and also statistically significant. Thus, we find a positive link between social activities and mental health among non-caregivers. We see that, when we measure caregiving and social capital roughly on the same scale, social ties are just as important as informal caring. A one-unit, or two standard-deviation, change in both variables leads to a similar response of mental health, albeit in different directions.

The parameter estimate for the interaction term equals 0.049 and is statistically significant, which provides evidence for the protecting role of social connections among caregivers. In other words, caregivers with a higher level of social capital have less mental health problems than their counterparts with less social capital.

To illustrate the interaction effect, it is useful to look at the predicted difference between non-caregivers and caregivers with respect to mental health for different values of the social capital index. This is shown in Figure (3.1) for an otherwise average person. Panel (a) shows the predicted mental health gap between non-caregivers and caregivers across increasing values of the continuous social capital index, based on the regression model using the standardized caregiver and social capital variables, respectively. One can clearly see that caregivers with more social capital suffer less, and that the mental health of caregivers and non-caregivers converges as the individual level of social capital increases. Consider, for example, an increase of the social capital index from one standard deviation below to one standard deviation above the mean, which compares low with high levels of social capital. The mental health difference between caregivers and non-caregivers decreases from -0.091 to -0.042. Thus, the fraction by which a one-unit increase in the social capital index reduces the negative association between caring and mental health amounts to 54%. In Panel (b), the mental health gap between non-carers and carers is estimated using a less parametric approach. That is, we run a regression of mental health on the caregiving indicator fully interacted with a categorical variable measuring the deciles of the social capital index. It can be seen that a social capital score above the fifth decile, or the sample median, basically renders the mental health difference insignificant.

The columns 2 to 6 of Table (3.2) show the regression results for the other mental health outcomes. We find a significant buffering role of social capital for the subscales depressive symptoms (MH) and perceived time pressure (TS). We also estimate a significant interaction between caregiving and social capital with respect to vitality (VT). Generally, the estimation results suggest that social capital may improve caregivers' mental health, even though it cannot fully compensate for the psychological cost of informal caring.

[14] We can calculate the association between the original caregiver indicator and mental health by dividing the parameter by two times the standard deviation of the original binary variable. Hence, caregiving reduces the mental health score by 0.268 ($0.066/(2 \cdot 0.123)$) which is equal to roughly 27% of a standard deviation.



(a) **(b)**

Source: Own calculations based on SOEP v30.

Figure 3.1: Difference in mental health between non-caregivers and caregivers, by different levels of social capital

3.4.2 The Role of Caring Intensity and Social Activities

This section offers a more detailed analysis of the moderating role of social capital in the context of informal care. The aim is to assess which caregivers derive the largest benefit from social capital on the one hand, and which kind of social activities are particularly helpful in ameliorating the adverse psychological implications of informal care on the other hand.

First, we assess whether the importance of social ties as an alleviating factor depends on the caregiving intensity. Social contacts might be particularly useful for caregivers with high time commitments. Panel A of Table (3.3) displays the estimation results from a regression model where we include the categorical caregiving intensity variable instead of the dichotomous indicator. It shows both the main terms and the interaction of the caregiving hours categories with the social capital index. The variables are again centered around the mean and standardized, so that the coefficients are comparable and correspond to a one-unit change in the explanatory variable. As expected, caregiving duration is negatively associated with the mental health score. The only significant (and positive) interaction with the social capital index is found for the highest category (35+ hours of care/week). This indicates that high-hour caregivers, who exhibit the highest psychological burden, might derive the greatest benefit from larger social networks. This result mirrors previous research by Cannuscio et al. (2004) who finds a similar pattern among female caregivers in a representative sample of registered nurses.

Second, specific social activities may matter more than others. To gain insights into the complex nature of social participation as a buffering factor, we substitute the individual social activity variables for the social capital summary index in the regression model. For each of the activities we create a dichotomous variable that indicates regular (at least monthly) participation. We obtain and include nine main terms reflecting frequent participation in social activities, and just as many interaction terms with the caregiver dummy variable. The results are shown in panel B of Table (3.3). The main term on the caregiver indicator now reflects the mental health difference between caregivers and non-caregivers among those individuals who rarely or

Table 3.3: The importance of caregiving hours and social activities

	Main terms		Interaction terms	
	Coef.	S.E.	Coef.	S.E.
(A) Measuring caregiving by hours of care per week				
1-9 hours	-0.023***	(0.008)	0.020	(0.016)
10-20 hours	-0.030***	(0.009)	0.008	(0.015)
21-34 hours	-0.044***	(0.011)	0.018	(0.019)
35+ hours	-0.041***	(0.012)	0.042**	(0.016)
$SOCI$	0.064***	(0.014)		
(B) Including social-activity variables separately				
Caregiver	-0.403***	(0.075)		
Social gatherings	0.048***	(0.012)	0.030	(0.083)
Helping	0.007	(0.009)	-0.015	(0.085)
Political participation	-0.004	(0.030)	-0.081	(0.284)
Religious participation	0.017	(0.017)	-0.084	(0.101)
Volunteer work	-0.004	(0.014)	0.300***	(0.091)
Sports participation	0.021*	(0.011)	0.092	(0.085)
Cultural attendance	0.010	(0.013)	0.088	(0.113)
Entertainment attendance	0.009	(0.012)	0.139	(0.123)
Artistic activities	-0.002	(0.014)	-0.037	(0.100)

$N = 70,680$. Robust standard errors, clustered at the individual level, in parentheses. Dependent variable: MCS = mental component summary scale. All regression models include controls for age (and age squared), marital status (married, separated, single, divorced, widowed), household size, number of children at ages 0-6 and 7-16, migration background, degree of urbanization, schooling (years of education), income (net household income), employment status (employed, not working, unemployed), working hours, hourly net wage, ownership status (owner vs. tenant), states, and survey years. Individual fixed effects are also included. In Panel A, the caregiving-hours indicators and the social capital are centered and standardized and have zero mean and a standard deviation of 0.5. The interaction terms contain the product between the rescaled hours-of-care and social-capital variables. In Panel B, the interaction terms consists of the caregiver dummy variable fully interacted with the social-participation indicators. Significance levels: *** p<0.01, ** p<0.05, * p<0.1.
Source: Own calculations based on SOEP v30.

never engage in social activities. Caregivers only seem to benefit from regular volunteer work, that is we exclusively find a positive interaction effect between caregiving status and unpaid public engagement. The implied buffering effect appears substantial. Regular voluntary work reduces the negative association between caring and mental health by 74%.

This result underlines the potential complementarity between caregiving and voluntary engagement. Caregivers may perform unpaid work in addition to providing informal care for a family member to distract themselves from the caregiving stress and burden. Further empirical support for this conjecture is provided by Burr et al. (2005) who have estimated a positive correlation between caregiving and voluntary work. This is consistent with the view that individuals contribute to charity to obtain psychological benefits. Unpaid work to help others provides a “warm glow” to the volunteer (see Andreoni, 1990), which may ease the caregiving taks or make the caregiving experience more pleasurable.

Our findings regarding volunteer work may provide evidence for the existence of a certain personality trait or motivation, that is some type of altruism, prompting individuals to assuming several helping tasks simultaneously. These individuals are often referred to as “super helpers”, who have a high willingness to help others and who may also derive the highest benefits from combining formal and informal caring activities (e.g. Hank and Stuck, 2008). However, only 14% of those who provide informal care also engage in regular voluntary work in our estimation sample. Whether the “super helper” personality is a widespread phenomenon can thus be disputed, at least when caring for ill family members is involved.

3.5 Sensitivity Analysis

In this section, we examine the sensitivity of the buffering effects caused by social capital. The results of several robustness checks are presented in Table (3.4). For comparison purposes, panel A repeats the baseline estimate for the moderator effect. Panel B and C examine whether observed characteristics, correlated with social capital, explain why informal carers with stronger social ties are less responsive to the caregiving burden. For instance, individual health status is one of the major correlates of caregiving status and social participation. It reflects a person’s capacity to provide informal care, and it may also influence how caregivers manage co-occuring duties and activities. Caregivers with larger social networks could therefore simply be healthier and more resilient than those who assume no caregiving tasks and have less social ties, which would explain stress buffering by social capital. Another important example is educational attainment. Better educated individuals may have greater economic, social, and psychological resources and knowledge that facilitate the caregiving duty (e.g. Huang et al., 2009; Ross and Wu, 1995; Stronks et al., 1998). Other factors such as age, marital status, the number of children, or the degree of urbanization are also potential moderators.

In panel B, we follow Dehejia et al. (2007) and include an interaction term between the caregiver dummy variable and a predicted social capital index. The latter is from a simple linear regression of the actual social capital index on all control variables, and thus represents a linear combination of all observed covariates. We are thus able to assess whether and to what extent the estimated interaction between caregiving and social capital changes when we control for the buffering effects of the remaining observed attributes. Panel C shows the results from an alternative procedure, using a matched sample of individuals where each respondent has roughly the same probability of reporting a high level of social capital. This is achieved by performing a nearest-neighbor matching procedure based on propensity scores. In the first observation period, individuals who ever report a $SOCI$ score above its median are matched with respondents who ever report a lower $SOCI$ score and who are closest in terms of observed covariates and propensity scores, respectively. The aim is to minimize the observed bias between these two groups. We therefore match to each high-$SOCI$ individual the five nearest neighbors within a caliper of 0.001. The caliper is a useful device to reduce the possibility of making poor matches (e.g. Morgan and Harding, 2006). This approach balances the distribution of observed covariates in these two groups. Hence, conducting a regression on this matched sample reduces concerns regarding the influence of observed characteristics, correlated with social capital, since they are basically the same for high-$SOCI$ and low-$SOCI$ individuals. In panel B and C, the coefficients for the interaction term between caregiving and social capital are practically equal to the baseline estimate. Thus, it is unlikely that the soothing influence of social capital is driven by observable factors correlated with social participation.

Table 3.4: The sensitivity of the moderator effect

	Main terms			
	Caregiver	SOCI	Interaction	Observations
(A) Baseline model				
	-0.066***	0.065***	0.049**	70,680
	(0.013)	(0.014)	(0.020)	
(B) Controlling for the buffering effect of observed caregiver characteristics				
	-0.067***	0.065***	0.050**	70,680
	(0.014)	(0.014)	(0.021)	
(C) Matched sample: High and low social capital				
	-0.078***	0.070***	0.051**	68,575
	(0.016)	(0.016)	(0.020)	
(D) Controlling for the buffering effect of caregiving context				
	-0.040	0.065***	0.045**	70,680
	(0.028)	(0.014)	(0.020)	
(E) Matched sample: Caregivers and non-caregivers				
	-0.058***	0.172****	0.026	7,710
	(0.013)	(0.065)	(0.022)	

Robust standard errors, clustered at the individual level, in parentheses. Dependent variable: MCS = mental component summary scale. The caregiver dummy variable and the social-capital index are centered and standardized and have zero mean and a standard deviation of 0.5. All regression models include controls for age (and age squared), marital status (married, separated, single, divorced, widowed), household size, number of children at ages 0-6 and 7-16, migration background, degree of urbanization, schooling (years of education), income (net household income), employment status (employed, not working, unemployed), working hours, hourly net wage, ownership status (owner vs. tenant), states, and survey years. Individual fixed effects are also included. The number of observations in the matched samples is smaller than in the original sample because some control individuals are not used in the matching procedure due to the common-support requirement. Significance levels: *** p<0.01, ** p<0.05, * p<0.1.

Source: Own calculations based on SOEP v30.

In panel D, we examine whether the buffering effect of social capital may also occur due to specific caregiving context characteristics. Individuals may benefit from a household member's social capital. Hence, persons with more social capital possibly provide informal care to individuals who are healthier per se. What is more, individuals who have stronger social ties are probably better able to acquire additional (private and public) support from sources outside the household. The relationship to the care recipient might also shape the experience of caregiving burden. It has been shown that individuals who provide care to spouses have an elevated risk of displaying mental health problems, because they more frequently lack social and other activities that could act as stress buffers (Pinquart and Sörensen, 2003).

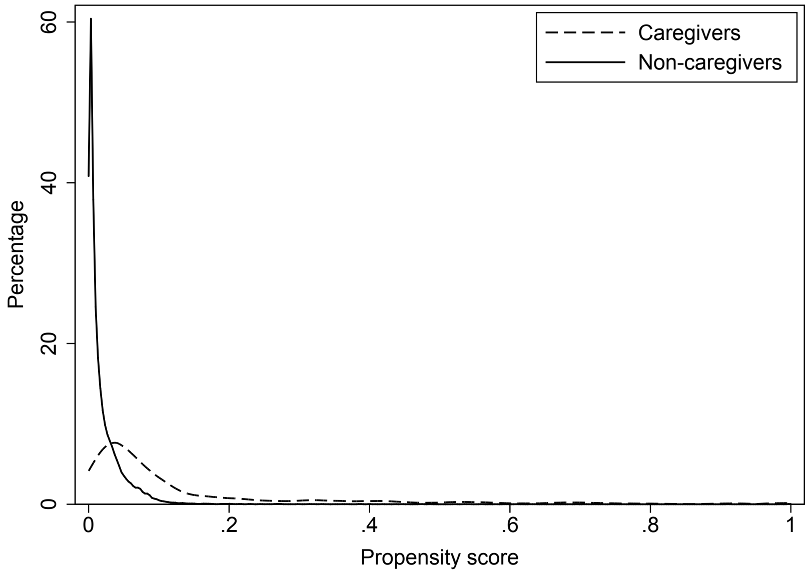
To capture the care-receiver's health status, we include a categorical variable that reflects the degree of limitations in daily self-care activities, as reported by the head of the household.[15] An alternative measure indicates to which of the three LTC levels the care-recipient had been assigned. This presumably is a more objective measure of the dependent's person health because the assessment is conducted by medical professionals. By entering the dependent's person health, we are also able to control for the so-called *family effect* of informal care (Bobinac et al., 2010). Having a family member with health problems may also adversely influence the psychological well-being of household members who are not the primary caregiver. Failure to control for these potential spillovers may lead to overestimates of the cargiving effect and thus may, *ceterus paribus*, underestimate the moderating role of social capital. What is more, we add a dummy variable which indicates whether the dependent person additionally receives help from other sources than the main caregiver in the household. Finally, to capture the relationship of the caregiver to the care receiver we construct a binary variable that equals one for spousal caregivers, and zero for all the other relations. For all variables, the reference group consists of individuals with the lowest category. As shown by the estimates in panel D, the interaction effect between caring and social capital is comparable to the baseline finding. This implies that the buffering effect of social capital cannot fully be explained by caregiving context.

Panel E deals with the problem that caregivers and non-caregivers are very different and that a fair comparison between carers and non-carers is complicated. Generally, the probability of providing informal care is very low in our estimation sample (1.5 percent), and there could be insufficient or poor overlap in terms of observed characteristics between carers and non-carers.[16] This means that for some values of the covariates there is no similar non-caregiver that can be compared with caregivers. This problem is illustrated in Figure (3.2). It depicts for both caregivers and non-caregivers the conditional probability of providing informal care and the propensity score, respectively. These quantities are equal to the predicted probabilities based on a probit model with caregiving as the binary outcome variable and observed characteristics as control variables. On the one hand, it shows that for most non-carers the theoretical probability of caring is very small or close to zero. On the other hand, the distribution of propensity scores is more widespread among caregivers. As a consequence, there is only a small range of propensity scores where there is sufficient overlap between carers and non-carers. Controlling for covariates in a linear regression may then produce biased estimates, because the linear model extrapolates the data obtained from overlapping covariate regions to make a comparison between carers and non-carers in the region with insufficient or even no overlap (e.g. Angrist and Pischke, 2008, p. 77).

To reduce the problems arising from poor overlap and extrapolation, we create a matched sample of individuals who have roughly the same probability of being or becoming a caregiver within the observation period. For this purpose, the nearest-neighbor matching procedure as described above is applied. We obtain a balanced sample where the observed covariate

[15] The questions on which this and the following variables are based, are described in the appendix in Table (3.A.3).

[16] Problems of insufficient overlap, or common support, are for example discussed by Smith and Todd (2005).



Source: Own calculations based on SOEP v30.

Figure 3.2: Distribution of the conditional probability of informal caring

distribution between carers and the control group are approximately the same. What is more, overlap, or common support, is ensured by restricting the estimation sample to individuals who have similar and positive propensity scores. The caliper of 0.001 also decreases the likelihood of using poor matches. The fixed effect regression is then performed using this homogeneous sample of individuals. As shown by panel E, using the matched sample and putting more weight on those who have roughly the same probability of becoming a caregiver, the point estimate of the interaction term almost halves compared to the baseline estimate. The fraction by which a one-unit increase in the social capital index reduces the negative association between caring and mental health decreases to roughly 37%. It should be noted, however, that enforcing the overlap requirement comes at the cost of losing a great deal of observations. This rises concerns that the obtained estimates may not be representative (e.g. Caliendo and Kopeinig, 2008).

3.6 Conclusion

This study examined the interrelationship between informal caring, social capital, and mental health. We tested the hypothesis that social capital weakens the negative association between informal care provision and mental health. To reduce concerns that unobserved heterogeneity might drive our results, we estimated fixed effects models which take unobservable, time-invariant differences between individuals into account. The results indicated that caregivers with more social capital, or more social ties, had better psychological well-being than caregivers with less social capital. Measuring caring and social capital essentially on the same scale, we found that a one-unit increase in the social capital variable reduces the negative association between caregiving and mental health by half. We also found significant moderating effects for vitality, depressive symptoms and perceived time stress. Further analyses revealed that particular groups of individuals might benefit more than others. Those with high caregiving workloads or caregivers who volunteer regularly appeared to experience the largest buffering effects.

We performed a variety of sensitivity checks to assess the robustness of our findings. Observed characteristics correlated with social capital could not explain the moderating role of social capital. Moreover, the buffering effect was unrelated to caregiving context, that is the care receiver's health status, support from individuals outside the household, and the relationship of the caregiver to the care-dependent person. Nevertheless, the buffering effect decreased substantially when the estimation sample was restricted to caregivers and non-caregivers with approximately the same probability of caring.

Our results may yield an explanation for the low utilization rates of formal caregiver support services. Some individuals (caregivers and care recipients) may prefer informal assistance obtained from family, friends, or neighbors to public or social programs. Clearly, policies to promote caregiver well-being should involve measures to facilitate information about the benefits of caregiver services and access to them. Yet, our findings also suggest that public programs should foster social and community involvement of caregivers (and care receivers).

To draw conclusive policy recommendations, however, further empirical evidence on the causal impact of social capital on the carergiver-health relationship is required. The results presented in this study rather represent associations, although we control for a variety of observed and unobserved background characteristics and alternative buffering mechanisms. It cannot be ruled out that unobserved shocks drive our findings. Another limitation results from the potential endogeneity between caregiving and social activities. Due to time restrictions, assuming caregiving tasks is usually associated with a decrease in other activities, such as leisure or social participation. To avoid that the potential simultaneity between caring and social activities influences the buffering effect, one would need external variation in both the decision to provide informal care and social participation. Generally, future studies should focus on causal buffering effects and the mechanisms that explain the moderating role of social capital.

Appendix

Table 3.A.1: Mental health scale (*MCS*) of the SOEP SF-12 survey

Subscale (No. of items)	Question wording and scales
	Scale: 1 (always) to 5 (never)
Vitality (1)	During the last four weeks, how often did you feel energetic?
Social functioning (1)	During the last four weeks, how often did you feel that due to physical and mental health problems your were limited socially, that is, in contact with friends, acquaintances, or relatives?
Role emotional (2)	During the last for weeks, how often did you feel that due to mental health or emotional problems...
Achieved less due to mental-health problems	...you achieved less than you wanted to at work or in everyday activities?
Less thorough due to mental-health problems	...you carried out your work or everyday tasks less thoroughly than usual?
Mental health (2)	During the last four weeks, how often did you...
Run-down, melancholy	...feel down and gloomy?
Well-balanced	...feel calm and relaxed?
Time stress (1) (not part of *MCS*)	Please think about the last four weeks. How often did it occur within this period of time, that you felt rushed or pressed for time?

Source: SOEP v30.

Table 3.A.2: Social activities in the SOEP study

Variable	Question wording and scales
	Which of the following activities do you take part in during your free time? Please check off how often you do each activity: at least once a week, at least once a month, less often, never.
Social gatherings	Meeting with friends, relatives or neighbors
Helping	Helping out friends, relatives or neighbors
Political participation	Involvement in citizen's group, political party, local government
Religious participation	Attending church, religious events
Volunteer work	Volunteer work in clubs or social services
Sports participation	Doing sports yourself
Cultural attendance	Going to cultural events (such as concerts, theatre, lectures, etc.)
Entertainment attendance	Going to the movies, pop music concerts, dancing, disco, sports events
Artistic activities	Artistic or musical activities (playing music/singing, dancing, acting, painting, photography)

Source: SOEP v30.

Table 3.A.3: Selected long-term care variables in the SOEP study

Presence of care-dependent person in HH: Does someone in your household need care or assistance on a constant basis due to age, sickness or medical treatment?

- *Yes/no*

Limitations in (I)ADLs: Who is that and by which of the following activities does he or she need assistance?

- *Errands outside of the house*
- *Running the household, preparing meals and drinks*
- *Minor care, such as help with dressing himself, washing up, combing hair, shaving*
- *Major care, such as getting in and out of bed, bowel movements*

LTC level: Does the person in need of care receive nursing care assistance?

- *Yes: Care level 1/Care level 2/Care level 3*
- *No*

Sources of help: From whom does this person receive the necessary assistance?

- *Relatives in the household*
- *Public or church nurse, social worker*
- *Private care service*
- *Friends*
- *Neighbors*
- *Relatives not in the household*

Source: SOEP v30.

CHAPTER 4

Health Satisfaction and Relative Health

4.1 Introduction

Relative concerns play a prominent role in human well-being and behavior. In sociology, relative concerns lie at the heart of the theory of relative deprivation. Well-being thus depends on the social context, and one feels relatively disadvantaged if one has less achieved than others in a social reference group (e.g. Stouffer et al., 1949; Runciman, 1966). Relative concerns are also of paramount importance in (social) psychological theories of social comparisons, according to which significant others act as a comparison standard for individual evaluations of one's own situation under uncertainty (Festinger, 1954).

In the economics literature, relative concerns have largely been analyzed in the context of consumption or income. The basic idea is that individual utility depends on both own income and the ratio or difference between own income and the average income of similar others. This connection has been termed the *relative income hypothesis* (Duesenberry, 1949; Easterlin, 1974). Relative income influences individual well-being because it provides important status effects. Income or consumption is thus seen as a positional good, whose value depends on the relative standing compared to significant others. As argued, for instance, by Solnick and Hemenway (1998), status concerned individuals prefer a situation where they have a relative advantage to a situation where they are absolutely better off. Consumption or income can thus create negative externalities which are not incorporated in individual decision making (Frank, 2008). Individuals who are concerned about their relative standing have an incentive to increase their consumption of the positional good, for instance through increased working hours. This leads to consumption levels that exceed the social optimum. The classical solution to the problem of negative externalities has been to increase taxes on the positional good (Frank, 1985).

This chapter takes up a different position and suggests that relative concerns with respect to health status are also relevant from an economic and public policy perspective. One reason as to why we should care about relative health effects is that many individuals evaluate their health status or changes in health relative to a reference point. When asked to judge their own health status, individuals tend to compare their situation with the condition of other people of the same age, gender, or who have similar health problems. This cognitive comparison process characterizes many situations in which survey respondents have to assess their health status based on condensed scales (e.g. bad, poor, satisfactory, good, very good), and it occurs even when the individual is not prompted to do so (Fienberg et al., 1985; Kaplan and Baron-Epel, 2003). As a consequence, subjective assessments of health status are prone to reporting bias. Persons with objectively the same health status may have different levels of self-rated health, or persons with different objective degrees of illness may have the same health perception (e.g. Groot, 2000). Thus, subjective health and quality of life is a relative concept, and respondents tend to evaluate their situation compared to that of stereotypical others.

Another reason as to why there should be primary interest in relative health effects are externalities. The canonical example for health-related external effects are positive physical health spillovers in the case of communicable diseases. For instance, the individual may directly benefit from the vaccination of others because the risk of infection decreases (Culyer, 1971).

External health benefits may also pertain to health status or health problems per se. Individuals with altruistic preferences derive utility from the good health status of other people, and they are willing to sacrifice own economic resources to contribute to other people's health and safety (e.g. Hurley and Mentzakis, 2013; Culyer and Simpson, 1980). However, analogous to the relative income literature, other people's health status may also produce negative externalities. Good relative health may provide important status effects. This hypothesis has been put forward by Mujcic and Frijters (2015) who assert that general health status is a positional good. The basic assumption is that good relative health provides economic advantages on marriage and labor markets in terms of better-off mating partners and more prestigious and better-paid jobs (e.g. Wilson, 2002; Hamermesh and Biddle, 1994). Thus, similar to relative income, positional externalities in the health domain increase the incentives to invest in own health in order to be able to keep up with others on labor and marriage markets.

The existence of relative health effects may have important implications for decision making in health care and for the design and funding of health care provision. First, significant health spillovers can lead to biased estimates of health status and quality of life. This becomes particularly relevant in economic evaluations of health policy programs. A critical role of other people's health for subjective health evaluations can bias the effectiveness of interventions. Consequently, cost-effectiveness ratios and thus rankings of alternative options are also biased, which may lead to decisions that entail a non-optimal allocation of resources (e.g. Groot, 2003; Labelle and Hurley, 1992).

The decisions regarding the design and financing of health care provision may also rest on the strength and direction of comparison health effects. For example, positive external effects due to altruistic preferences indicate that individuals may have a preference for public provision of health care. This is consistent with the view that there are positive spillovers in the context of health status. Generally, altruistic and rational individuals are willing to spend own financial resources to facilitate access to health care among disadvantaged individuals. The existence of such externalities justifies the public and widespread provision of health care services (e.g. Culyer and Simpson, 1980). In a solidarity-based health insurance scheme, low-risk individuals are willing to pay a relatively high premium to subsidize the cost of health care among high-risk individuals (e.g. Paolucci, 2011).

Given that individual health status or health care consumption reduces other's utility or value of health, there are strong private incentives to invest in health. Analogous to Frank (2008), since good health may provide economic status effects, everybody increases the consumption of medical goods and services. This may not necessarily change the relative position in the societal health distribution, but can lead to an excessive demand for medical care. This problem is potentially aggravated in social health insurance, which effectively decreases the cost of medical care for the disadvantaged. Thus, the private costs are lower than the social cost, which may induce welfare losses due to moral hazard. Thus, as argued by Paolucci (2011), positional externalities in the health domain imply that decisions regarding the funding and design of health care provision need to balance the positive and negative welfare effects of increased health care consumption. For instance, a response to excessive demand for medical care or moral hazard could be to reduce insurance coverage for goods and services where positional concerns dominate.

This chapter provides an empirical test of the relationship between relative health status and individual health satisfaction. For this purpose, we rely on longitudinal individual-level data from the German Socio-Economic Panel (SOEP) study. Alongside a large set of demographic and socioeconomic variables, this survey retrieves information on health-related quality of life and subjective well-being for a number of years. We use the health satisfaction scale to measure the individuals subjective evaluation of health status. It can be associated with the utility derived from different health states and provides a practical alternative to quality-adjusted life years (QALYs) to assess the benefits or losses of health care interventions (e.g. Cutler and Richardson, 1997; Van Praag and Ferrer-i-Carbonell, 2008). Moreover, Graham (2009) and Frijters et al. (2011) have shown that health is an important predictor of mortality.

To measure individual health status, we calculate an overall health index based on two summary scales that capture the respondent's physical and mental illness. Other people's health is calculated based on the average health index in the individual's reference group, which consists of persons with similar demographic and socioeconomic characteristics. We employ linear panel data models with health satisfaction as the dependent variable and own health, reference-group health, and a variety of observable characteristics as explanatory variables. Since respondents are followed over several waves, we can apply the fixed effects estimator to eliminate unobserved heterogeneity correlated with both the outcome variable and regressors and which is constant over time.

We test whether and how changes in reference-group health, given own health, influences individual health satisfaction. We also divide the sample into those who are healthier and sicker than their reference group, respectively. Social psychological research has shown that it matters for the effect of social comparisons whether one is below or above the comparison standard (e.g. Buunk et al., 1990), although the consequences are ambiguous. Another specification includes an interaction term between the respondent's health index and the average health index of the reference group. Reference-group health may alter the relationship between own health and health satisfaction. The subjective evaluation of own ailments therefore depends on whether they are widespread in the social reference group, or the norm to be healthy is weak (e.g. Powdthavee, 2009).

Generally, we do not find strong effects of relative health on individual health satisfaction. Changes in reference-group health are basically unrelated to changes in satisfaction with health. Only the interaction term between own health and reference-group health is significant, which would suggest that social health norms could lead to biased satisfaction ratings regarding own health status. However, the estimated association is quantitatively and economically negligible. All in all, the empirical results provide evidence that relative health concerns are rather unimportant for individual health evaluations, which is in accordance with the results presented by Powdthavee (2009). Thus, the bias in cost-effectiveness analyses due to reference-group effects seems to be rather marginal. What is more, the results indicate that health-related externalities should not strongly influence the decision on how to design and finance health care provision.

The remainder of this chapter is organized as follows: Section (4.2) reviews the literature on the role of other people's health problems for own individual health and happiness. Section (3) discusses the empirical approach employed in this study. It includes a description of the dataset and variables, followed by an exposition of the empirical method. Section (5) assesses the sensitivity of the results. Section (6) summarizes the main findings, outlines limitations, and offers future research avenues.

4.2 Related Literature

4.2.1 Status Concerns and Altruism

Theoretically, the effect of relative health is ambiguous. A negative association between reference-group health and individual health satisfaction may occur due to status concerns. Besides psychological benefits, a relative advantage in the health domain provides tangible economic benefits, for instance on marriage and labor markets (Mujcic and Frijters, 2015). Health investments are seen as a critical precondition for marital success, and healthier individuals tend to find superior mating partners in terms of health and socioeconomic status (e.g. Wilson, 2002). Furthermore, health status also determines labor market achievement. For example, it has been found that good health status, and also physical fitness and attractiveness, is positively related with the chance of having a prestigious job and high earnings (e.g. Hamermesh and Biddle, 1994; Lindeboom, 2006). In this sense, health status is a positional good and produces negative externalities for those who are less healthy. An explanation of positional concerns in the health domain are preferences related to envy. Comparing with others who are or become healthier may simply reduce satisfaction with own health status, because one lacks what seems to be a socially accepted standard. A natural response would be, for instance, to emulate other people's health (e.g. Elster, 1991). Whether someone attains a higher social position not only depends on health status per se, but also on how own health, fitness, or physical appearance compares to others who compete for the same partner or job. Following Frank (2008), when there are others who are healthier, individuals have an incentive to invest in their own health and increase health-improving behavior, respectively. All in all, if individuals are status-concerned, an increase in reference-group health decreases individual health satisfaction, since the value of own health diminishes.

Empirical evidence for positional concerns in the health domain is provided by Mujcic and Frijters (2015), who find a negative association between reference-group health status and life satisfaction using Australian survey data. Furthermore, status effects appear to be important with respect to one's physical appearance. As shown by several authors, individuals care about their position in the social weight and height distribution, and being slimmer or taller than peers can increase well-being (e.g. Blanchflower et al., 2009; Carrieri and De Paola, 2012; Oswald and Powdthavee, 2007). In contrast, there is rather direct evidence based on preference-elicitation methods showing that relative concerns in the health domain are rather negligible (e.g. Alpizar et al., 2005; Carlsson et al., 2007; Hillesheim and Mechtel, 2013; Solnick and Hemenway, 2005).

A positive association between reference-group health and individual health satisfaction can be expected when there are positive health spillovers. External health benefits arise because many individuals care about other people's health status (Culyer and Simpson, 1980). This leads to interdependent preferences where the health of others positively influences subjective well-being. Hence, due to feelings of sympathy, individuals may evaluate their health status more favorably if they observe that the health of peers improves. This has been shown by Carrieri (2012), who finds a negative association between the health problems of peers on the one hand and happiness and self-assessed health on the other hand. However, it could be argued that altruistic preferences or sympathy are less important when individuals compare with individuals from their larger social network. Altruism is thus likely to be more relevant towards others in a person's immediate environment, such as the family or relatives (e.g. Groot and Van Den Brink, 2003; Viscusi et al., 1988).

4.2.2 Upward and Downward Comparisons

The comparison health effect may depend on whether the individual is healthier or sicker than the reference group. Useful guidance on the multifaceted consequences of social comparisons comes from social psychological research. Social comparison theory basically distinguishes between comparisons with better-off (upward comparisons) and worse-off (downward comparisons) individuals, both of which may have different effects on subjective well-being and health evaluation.

Clearly, positional or status concerns in the health domain can lead to negative effects of upward comparisons. Being sicker than someone with a similar background may create distress and reduce the ability to compete on marriage and labor markets. We may also call this a situation of relative deprivation in terms of health status, because the individuals falls short of the prevailing health standard of his or her reference group (see also Runciman, 1966; Yitzhaki, 1979). Accordingly, an improvement of reference-group health decreases health satisfaction for sicker individuals.

Nevertheless, it is also possible that comparisons with healthier individuals actually increase health satisfaction. Relatively disadvantaged individuals may appreciate the information provided by this situation, namely that their health status can improve in the future. Healthier individuals can provide valuable information that may assist in problem-solving (e.g. Buunk et al., 1990). Hence, a positive association between reference-group health and health satisfaction among sicker individuals can be expected when the information effect predominates.

The relative health effects may also differ among healthier individuals. Observing that the reference group is worse-off in terms of medical condition may increase health satisfaction. Status concerns in the context of health would lead to positive effects of downward comparisons. As discussed above, one could obtain a higher social rank from being healthier than similar others, which presumably increases the value of one's own health. In line with the income comparison literature, we can term this situation relative satisfaction (Wunder, 2009; Yitzhaki, 1979), because the individual is above the socially determined health standard. This is also in line with social comparison theory, which basically states that downward comparisons primarily serve to enhance mental well-being (Wills, 1981). Hence, an improvement of reference-group health would reduce the relative advantage, and therefore decrease satisfaction with health among healthier individuals.

Nonetheless, downward comparisons can also have negative consequences for individual health evaluation. Altruistic concerns or feelings of sympathy may produce negative psychological consequences from observing that others are worse-off. Individuals might thus evaluate their health less favorably when comparing with sicker peers. Negative effects of downward comparisons can also be justified based on preferences for equitable outcomes. Some individuals may simply dislike inequalities, be it to their advantage or their disadvantage (Fehr and Schmidt, 1999). Thus, if healthier individuals are altruistic or inequity averse, better reference-group health increases individual health satisfaction.

4.2.3 Social Norms

Reference-group health could influence individual health satisfaction also via altering the effect of own health status. Under certain circumstances, the negative effect of own illness on health satisfaction reduces when the health of peers worsens. The perceived health status of significant others can act as a social norm or standard, according to which individuals evaluate their own situation. From an economic perspective, norm-guided behavior may arise from rational considerations. On the one hand, the violation of social norms, such as overeating, may have devastating psychological effects such as feelings of embarrassment, anxiety, guilt or shame (Elster, 1989). On the other hand, individuals have an incentive to follow social norms because they fear the threat of social sanctions, and want to avoid other people's disapproval.

The effects of social health norms on individual health satisfaction can be analyzed using the theory of social customs (Akerlof, 1980). Originally applied to the causes of unemployment persistence, it can also yield important insights into the relationship between health evaluation and health-related social standards. The theory of social customs has been applied to the health domain by Powdthavee (2009). Hence, health satisfaction ratings are influenced by the individual's reputation within the reference group, which is in turn a function of the propensity to adopt the norm and the share of individuals following norm. Theory predicts that if the perceived norm to be healthy weakens, for instance if the health status in the reference group worsens, the negative health satisfaction effect of own illness decreases. Put differently, the health-satisfaction gap between healthy and sick individuals decreases as the average health of the reference group diminishes. Powdthavee (2009) also provides empirical evidence for health norm effect. He shows that an increase in other household member's health problems alleviates the negative relationship of own diseases with subjective health. Conversely, increasing levels of reference-group health should generally improve individual health satisfaction, because it enhances the positive effect of good individual health status.

4.2.4 Hypotheses

Table (4.1) summarizes the expected relationships between health comparisons and satisfaction with health. As the literature review has shown, the mechanisms might differ according to whether individuals are healthier or sicker than their reference group. We therefore distinguish between general associations on the one hand and the potential reference-group effects on individual health satisfaction among those how are healthier or sicker on the other hand. Panel A illustrates the general influence of reference-group health on individual health satisfaction. When status concerns dominate, improvements in reference-group health *ceterus paribus* worsen the individual's relative position and thus lower health satisfaction or the value of own health. The health of the reference-group may also influence health satisfaction by altering the effect of own health (problems) on health satisfaction. If reference-group health, or the perceived social health norm, increases, the psychological cost of not satisfying the norm rises. Hence, we hypothesize that reference-group health rises individual health satisfaction by improving the psychological benefits of own good health. When individuals have altruistic preferences and care about other people's health, increasing reference-group health should raise individual health satisfaction.

Table 4.1: Associations between reference-group health and health satisfaction

Effect on health satisfaction	Mechanism
A: General associations/all individuals	
−	Status concerns, worsening of relative position/advantage
+	Social health norms, psychological gain of good health status larger when the health norm increases
+	Altruism, individuals care about other people's health
B: Healthier individuals	
−	Status concerns, worsening of relative position/advantage
+	Altruism, individuals care about other people's health; inequity aversion
C: Sicker individuals	
−	Status concerns, worsening of relative position
+	Information, signal

Source: Own illustration.

Panel B shows the hypothesized effects of reference-group health among those who are healthier than the comparison standard. Status preferences in the health domain imply a negative relationship between reference-group health and individual health satisfaction. When the health of significant others improves, the relative advantage and thus the ability to attain a higher social position worsens, which presumably reduces the value of own health. However, healthier individuals who dislike inequalities or have preferences based on altruism may approve of the health improvements among the worse-off. In this case, the association between reference-group health and health satisfaction is positive among healthier individuals.

Panel C demonstrates how reference-group health might affect satisfaction with health among those who are sicker than their comparison group. Having less than similar others may result in feelings of relative deprivation. If health is a positional good or individuals are status concerned, rising levels of reference-group health could amplify the psychological consequences of relative deprivation. Against this background, the association between reference-group health and individual health satisfaction is negative. Nevertheless, observing that the health of significant others improves can provide important information for the worse-off. They could interpret it as a signal that there are opportunities to improve health status and to catch up, respectively. If sicker individuals focus on these positive aspects, the effect of reference-group health on their health satisfaction is positive.

4.3 Empirical Approach

4.3.1 Data and Variables

This study uses longitudinal data from the German Socio-Economic Panel (SOEP) (Wagner et al., 2007).[17] The SOEP study is a representative survey of about 11,000 households located in the Federal Republic of Germany and is particularly useful for our purposes, since it provides information on the individuals' health perceptions as well as rather objective health indicators for a number of years. The baseline estimation sample consists of an unbalanced panel of 25,788 individuals over six survey years, leading to 93,674 person-year observations.

Dependent Variable

The dependent variable in this study is the satisfaction-with-health scale. In the SOEP, respondents are asked to answer the following question on a scale ranging from 0 ("totally unsatisfied") to 10 ("totally satisfied"):

How satisfied are you with your health?

Health satisfaction scores provide a practical alternative to health-utility elicitation methods in large-scale samples. According to Van Praag and Ferrer-i-Carbonell (2008), answers to the health satisfaction question can be associated with the utility derived from different health states. It is therefore strongly related to the QALY approach to assess the benefits or losses of health care interventions. A similar approach has been adopted by, for example, Cutler and Richardson (1997) and Groot (2000) to estimate the impact of illnesses on subjective health. They also show how the estimated regression coefficients of the disease variables can be used to calculate QALY weights. Moreover, health satisfaction seems to be an important indicator for the general health of the population, because previous literature shows that satisfaction with health is an important and independent predictor of mortality (Frijters et al., 2011; Graham, 2009).

The role of relative health concerns has largely been neglected in the domain of health satisfaction. To the best of our knowledge, there is only one study by Winkelmann and Studer

[17]Socio-Economic Panel (SOEP), data for years 1984-2012, version 29, SOEP, 2013, doi:10.5684/soep.v29.

(forthcoming) who relate a measure of relative physical illness to individual health satisfaction scores using the SOEP data. However, they only rely on a cross-section of individuals and do not consider other mechanisms that could justify the importance of other people's health problems.

Own Health and Relative Health

To measure individual health status, we rely on generic measures of physical and mental well-being, which are included every two years in the SOEP questionnaire since 2002. The short-form 12 questionnaire (henceforth SF-12) is a brief version of the SF-36 questionnaire, and is a widely accepted and validated tool for the measurement of health-related quality of life (e.g., Andersen et al., 2007). It consists of twelve self-reported items that comprehensively measure the respondents' physical and psychological health. These items are merged into eight subscales and summarized into two aggregate dimensions via exploratory factor analysis: "physical health" (*pcs*) and "mental health" (*mcs*).[18]

Following Frick and Ziebarth (2013), we use the average of the *pcs* and the *mcs* scores to obtain a measure of overall health. For the ease of interpretation, this measure is standardized with mean zero and a standard deviation of one (z-standardization). It has been shown that the health assessment based on the SF-12 provides more objective health measures than single self-assessed health scales. For example, Ziebarth (2010) shows that the physical component of the SF-12 in the SOEP does not suffer from reporting heterogeneity related to socioeconomic status. Moreover, this measure comprehensively captures the individual's health status, summarizing physical and mental-health aspects. Previous studies, in contrast, have used a checklist of several diseases or diagnoses as objective health indicators (Carrieri, 2012; Powdthavee, 2009). However, these measures rather reflect the existence of specific conditions and leave out many aspects of physical and mental health (see also Eibich and Ziebarth, 2014). Finally, strictly objective health indicators, like grip strength, are only occasionally assessed in large-scale surveys like the SOEP.

Relative health is defined as the average health index of all persons in the respondent's reference group.[19] This measure is also z-standardized. An interaction term between the respondent's own health index and the average health index of the reference group is constructed to test for social health norm effects (see also Carrieri, 2012; Powdthavee, 2009).

Control Variables

We include additional regressors that are potentially important for health satisfaction, health status, and social comparisons. These are age, age squared, sex, equivalized net household income, years of education, number of adults and children in the household, employment status *(employed, not employed, unemployed)*, marital status *(married, single, widowed, divorced, sep-*

[18]See Table (4.A.1) in the appendix for a detailed description of question wording and response scales of the SF-12 in the SOEP.

[19]Note that the individual's contribution is excluded from the calculation of the average health index.

Table 4.2: Summary of variables

Variable	Mean	Std. Dev.	Min	Max
Health satisfaction	6.55	2.19	0	10
Own health	0	1	-3.85	1.70
Ref.-group health	0	1	-4.08	3.09
Female	0.52	0.50	0	1
Age	50.32	16.27	17	102
Net household income	21,992.59	11,658.72	2,815	136,554
Years of education	12.23	2.70	7	18
No. of adults	2.16	0.83	1	9
No. of children	0.48	0.86	0	9
Employed	0.59	0.49	0	1
Not employed	0.36	0.48	0	1
Unemployed	0.06	0.24	0	1
Married	0.64	0.48	0	1
Single	0.19	0.40	0	1
Widowed	0.07	0.25	0	1
Divorced	0.08	0.27	0	1
Separated	0.02	0.13	0	1
East German	0.27	0.45	0	1

Note: $N = 93,674$.
Source: Own calculations based on SOEP v29.

arated), and dummy variables for East Germans[20], nationality (non-German) and survey year. All continuous variables are included in their logarithmic form. Table (4.2) summarizes the variables used in the analyses.

Definition of the Reference Group

One practical issue pertains to the construction of the reference group. It is well understood by social psychologists that individuals strategically choose their comparison targets (e.g., Buunk and Gibbons, 1997). They could, for example, intentionally seek contact with better-off or worse-off individuals to gain valuable information and to increase well-being, respectively. Thus, specific characteristics of the individual (for example health status or ability) likely determine with whom he or she compares, and how this comparison affects his or her well-being. From an economic perspective, this creates a problem of endogeneity (e.g., Falk and Knell, 2004). There is no information in the general SOEP questionnaire on the relevant reference groups and comparison standards, and we assume that the reference group is exogenously given. Clearly, we never know the true reference point of individuals and therefore measure the respondent's reference-group with error.

Some evidence on which reference groups are used when people compare their health status is provided by Kaplan and Baron-Epel (2003). They identify persons of the same age and friends and acquaintances as the most influential reference groups in the subjective evaluation of health.

[20]The dummy variable equals one in case the respondent had lived in East Germany in 1989.

Based on this information, we define a synthetic reference group characterized by a combination of demographic and socioeconomic attributes (see also Ferrer-i-Carbonell, 2005). We assume that the respondents compare their health with people of the same sex, age, educational level and living in the same region. We construct seven age groups ($< 25, 25-34, 35-44, 45-54, 55-64, 65-74, > 75$), four education groups (*dropout/basic track, intermediate track, academic track, other*), and 16 regional groups based on states ("Bundesländer"). The focus of the analysis at hand is on contemporaneous, interpersonal health comparisons. Hence, the reference groups are created on an annual basis. The analysis is restricted to reference groups with at least ten individuals, to avoid that the calculations are based on a (very) small number of observations. The average group size amounts to 62 observations, ranging from a minimum of 10 to a maximum of 222 individuals.

In the absence of information on the true comparison standard, we believe that the peer-group specification adopted in this study appears to be a good approximation. It most likely reflects the individual's social network and the set of persons with whom individuals frequently interact. What is more, it captures the most important reference groups mentioned by Kaplan and Baron-Epel (2003), which consists of friends and acquaintances of the same age. We additionally include gender and educational attainment to be better able to capture the respondent's social ties. However, in a robustness check, we also add the respondent's occupational level as a reference-group characteristic. This is motivated by the conjecture that (working) individuals most of the time interact with their colleagues.

4.3.2 Estimation Method

To test whether the level of reference-group health predicts individual health satisfaction, the following empirical model is specified:

$$S_{it}^{*} = \alpha H_{it} + \beta H_{it}^{r} + \delta' X_{it} + u_i + \varepsilon_{it}. \tag{4.1}$$

Equation (4.1) includes the individual's own health index (H_{it}) and the average health index of individual i's reference group (H_{it}^{r}). The dependent variable S_{it} measures individual $i's$ health satisfaction at time t. As discussed above, the coefficient β can be positive or negative depending on whether relative health confers status effects or individuals are altruistic. This specification assumes that the effect of reference-group health is the same for all individuals. However, we also estimate this model for healthier and sicker sub-samples to test whether the effects of reference-group health differs across individuals who are healthier and sicker than their reference group, respectively.

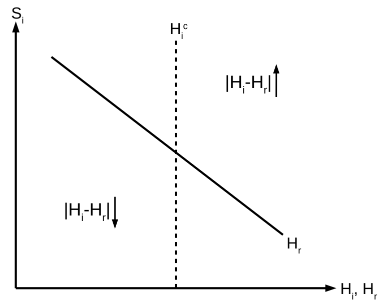
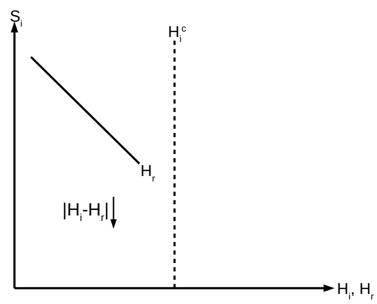
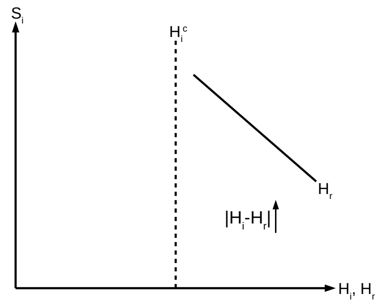
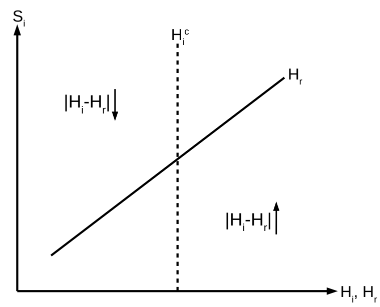
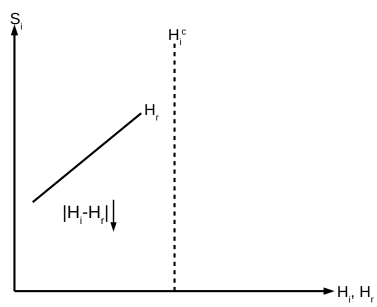
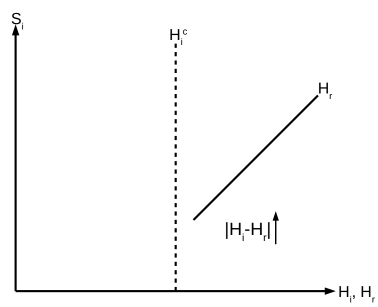
Following Carrieri (2012) and Powdthavee (2009), the second specification adds an interaction term between the individual's own health index and the reference-group health index to assess the influence of health-related social norms, that is

$$S_{it}^{*} = \alpha H_{it} + \beta H_{it}^{r} + \gamma H_{it} \times H_{it}^{r} + \delta' X_{it} + u_i + \varepsilon_{it}. \tag{4.2}$$

We expect that the coefficent on the interaction term, γ, is positive. This implies that increasing levels of reference-group health favor good health status and aggravate the negative psychological consequences of health problems, respectively.

The parameter β shows how variations of reference-group health given own health status influence individual health satisfaction. An equivalent approach allows us to interpret the estimates as changes in the relative difference between own health status and reference-group health. Table (4.3) illustrates the stylized relationships between reference-group health, the difference between own health and reference group health, and individual health satisfaction. For the sake of simplicity, we suppress the time subscript and denote individual health status as H_i, the average health in the individual's reference group as H_r, and individual health satisfaction as S_i. The graphs basically show how variations in reference-group health H_r change the difference $H_i - H_r$ when own health status is fixed at some value H_i^c. The vertical axis shows the range of values of health satisfaction and the horizontal axis measures H_i and H_r, respectively.

Table 4.3: Illustration of reference-group effects

All individuals (1)	Healthier Individuals (2)	Sicker individuals (3)
(a) Status concerns	*(a) Status concerns*	*(a) Status concerns*
		
$H_r \uparrow \Rightarrow S_i \downarrow$	$H_r \uparrow \Rightarrow S_i \downarrow$	$H_r \uparrow \Rightarrow S_i \downarrow$
(b) Altruism	*(b) Altruism*	*(b) Information*
		
$H_r \uparrow \Rightarrow S_i \uparrow$	$H_r \uparrow \Rightarrow S_i \uparrow$	$H_r \uparrow \Rightarrow S_i \uparrow$

Source: Own illustration.

The first column of Table (4.3) lays out the influence of reference-group health for all individuals. Therefore, the difference $H_i^c - H_r$ can be positive or negative. If $H_i^c > H_r$, increasing values of H_r reduce the health status difference between the individual and the reference-group.

If, in contrast, $H_i^c < H_r$, increasing values of H_r increase the distance between individual health and other people's health in absolute terms. The effect of H_r on S_i depends on whether status concerns or altruistic preferences dominate. If individuals focus on the status effects of health (panel (a)), improving levels of reference-group health lead to a deterioration of the individual's relative position which should decrease health satisfaction. However, if individuals have preferences based on altruism (panel (b)), any improvement in other people's health is appreciated and better reference-group health is associated with higher health satisfaction.

The second column of Table (4.3) refers to the effects of reference-group health among those who are healthier than similar others. That is, H_i^c is always greater than H_r and the difference $H_i^c - H_r$ is strictly positive. Hence, increasing values of H_r unambiguously decrease the health difference between the individual and his or her reference group. As shown in panel (a) of the second column, status-concerned individuals would interpret this as a worsening of their relative advantage and improving levels of reference-group health reduce the value of their own health and health satisfaction, respectively. Nevertheless, if individuals care about other people's health or dislike inequalities per se, as illustrated in panel (b), observing that worse-off peers become healthier can increase own health satisfaction.

The third column of Table (4.3) shows how reference-group health influences individual health satisfaction among those who are sicker than their reference standard. The difference $H_i^c - H_r$ is therefore always negative, and increasing values of reference-group health actually widen the distance between individual health status and reference-group health in absolute terms. On the one hand, when status concerns dominate (panel (a)), improving reference-group health is associated with lower levels of health satisfaction. The health gains of the better-off reinforce the relative disadvantage of sicker individuals which can create psychological distress or stronger feelings of relative deprivation. On the other hand, the effect on health satisfaction can be positive if individuals value the informational or signaling effects of other people's health improvements in their social environment (panel (b)).

We use the so-called probit-adapted ordinary least squares (POLS) approach developed by Van Praag and Ferrer-i-Carbonell (2008) to estimate the regression models. Clearly, the ordinal nature of the dependent variable requires the application of an ordered response model. However, the estimation of ordinal regressions is often associated with computational problems, and the interpretation of interaction effects is generally cumbersome in non-linear regression models (Ai and Norton, 2003; Greene and Hensher, 2010; Mallick, 2009). The ordinal dependent variable is thus roughly cardinalized, by calculating the conditional expectation based on a standard normal distribution:

$$S_{it}^c = E\left(S_{it}^* | \mu_{J-1} < S_{it}^* \leq \mu_J\right) = \frac{\phi(\mu_{J_{z-1}}) - \phi(\mu_{J_z})}{\Phi(\mu_{J_z}) - \Phi(\mu_{J_{z-1}})}, \tag{4.3}$$

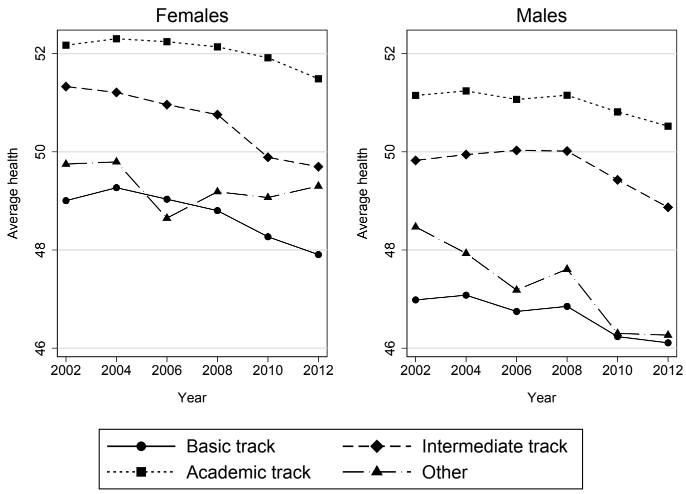
where S_{it}^c is the cardinalized health satisfaction variable, Φ is the standard normal cumulative distribution, and ϕ is the density function of the standard normal distribution. The cut-off points μ_{J_z} are retrieved by calculating the z-values of the cumulative frequencies that correspond to each value of the ordinal health satisfaction scale. This procedure leads to a transformed health satisfaction variable that is no longer bound between 0 and 10.

We replace S_{it} by the transformed variable S^c_{it} in the specifications presented above, enabling us to use linear panel-data methods. The coefficients on explanatory variables are thus directly comparable with the results of an ordered probit model except for a multiplication factor. Moreover, since the transformed dependent variable is approximately standard normal, the estimated coefficients on the independent variables can be interpreted as standard-deviation changes in the predicted "quasi-cardinal" health satisfaction scores.

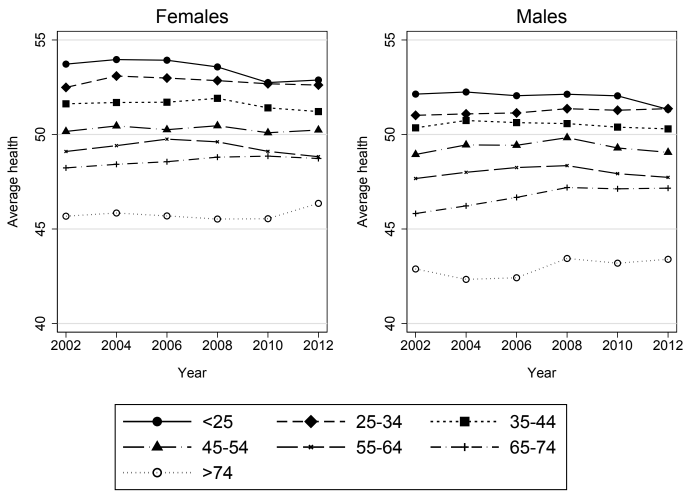
We estimate fixed effect models to take unobserved heterogeneity into account. There are many characteristics like personality, cognitive skills or motivation that simultaneously influence health satisfaction, individual health status and reference-group health, but are usually unobserved in observational studies like the SOEP. Likewise, personality traits, such as the tendency to socialize or emotional instability, may heavily influence both the propensity to compare with others and the affective consequences of social comparisons (e.g. Gibbons and Buunk, 1999). The fixed effects model reduces concerns that unobserved factors bias the empirical estimates. It eliminates confounding due to time-invariant unobserved factors by relying on within-individual variation in the dependent and independent variable. In fact, this cancels out any variable, observed and unobserved, that is stable over time within respondents. This approach is equivalent to control for individual effects by including dummy variables for each respondent. As a result, the empirical findings provide unbiased estimates as long as unobserved confounding due to time-varying characteristics is unimportant. Furthermore, we assume that the potential simultaneity between individual health satisfaction, own health and peer-group health does not bias our results. A potential drawback of the fixed effects model is that it may suffer from attenuation bias. It may well be that the variation of reference-group health over time is rather low, which would drive the fixed effects regression coefficients towards zero.

The fixed effects estimates measure how changes in reference-group health, or the relative health difference as illustrated in Table (4.3), influence changes in health satisfaction. To be more precise, the coefficients measure the influence of transitory changes in reference-group health or the relative health distance. Changes of comparison health may come from both variations of average health status within and across reference groups. We argue that these are largely driven by individuals getting older. This seems to be a reasonable approach assuming that the other reference-group attributes gender and education are (practically) fixed, and that frequent migration between states or over longer distances in Germany is negligible (e.g. Jürges, 1998; Pfaff, 2012).

As an example, Figure (4.1) shows that the variation in the health status of others combines both changes within reference groups and across different reference groups, based on selected attributes that define an individual's comparison group (sex, education, and age). Figure (4.1a) depicts the health trend across the different educational categories. The graph shows that within reference groups defined by educational level, which is rather time-constant, the average health status deteriorates over time. The evolution of health within reference-groups is likely to be age-related, since the reference groups are defined over a five-year age interval. Figure (4.1b) demonstrates that at any given point in time, older reference groups have worse health than younger reference groups. Since individuals are assumed to compare their health status with



(a) Educational level



(b) Age group

Source: Own calculations based on SOEP v29.

Figure 4.1: Health status and reference-group attributes

others of similar age, they also adjust their comparison group over the life cycle. This implies that at some point in time they change their reference group which likely leads to a variation in the comparison health level.

4.4 Results

4.4.1 Main Results

Before proceeding to the fixed effects results, we present estimates using the pooled OLS estimator. Table (4.4) shows that own health is strongly and positively associated with health satisfaction. For instance, as shown in column 1, a one-standard-deviation increase in the individual's health index roughly raises the health satisfaction score by 0.6 standard deviations

Table 4.4: Health satisfaction and relative health, OLS results

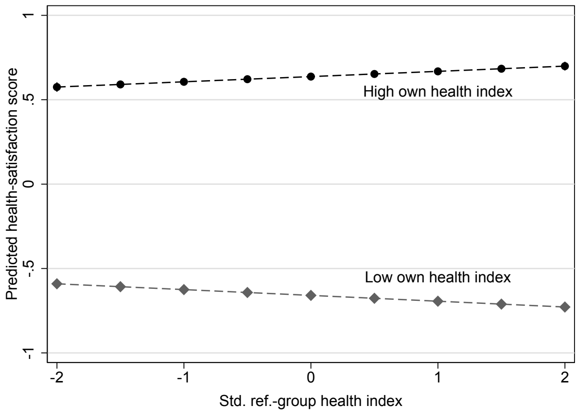
	All individuals		Healthier	Sicker
	(1)	(2)	(3)	(4)
Own health	0.641***	0.649***	0.789***	0.589***
	(0.003)	(0.003)	(0.009)	(0.006)
Ref.-group health/10	-0.073	0.004	-0.235***	-0.167**
	(0.054)	(0.055)	(0.073)	(0.080)
Own health ×		0.034***		
Ref.-group health		(0.003)		
Control variables	Yes	Yes	Yes	Yes
Individual FE	No	No	No	No
No. of observations	93,674	93,674	52,033	41,641
No. of individuals	25,788	25,788	19,068	17,116

Note: Robust standard errors, clustered at the individual level, in parentheses. *Health satisfaction* is measured with the quasi-cardinal and approximately standardized health-satisfaction scores. *Own health index* and *reference-group health index* are standardized such that their mean equals zero and their standard deviation equals one. Control variables include sex, age, net household income, years of education, number of adults in the household, number of children in the household, employment status, marital status, a dummy variable for East Germans, and dummy variables for 16 states and 6 survey years. Significance levels: *** p<0.01, ** p<0.05, * p<0.1.

Source: Own calculations based on SOEP v29.

in the full sample. This relationship is significant at the 99 % level. Across all specifications and models, reference-group health relates negatively to individual health satisfaction, which provides some evidence for status concerns in the health domain. The coefficients measure the change in health satisfaction when the reference-group health index changes by 10 standard deviations. Thus, the estimated associations are rather small compared to the effects of own health status. For instance, the columns 3 and 4 illustrate that an increase of reference-group health by one standard deviation significantly reduces health satisfaction by roughly 0.02 standard deviations.

We also find significant health norm effects, as indicated by the significant interaction term between own health and reference-group health in the second column. The positive sign of the interaction term suggests that a higher social health standard favors individual health improvements. Conversely, increasing levels of reference-group health aggravate the adverse effect of declining health status on individual health satisfaction. The health norm effect based on the estimates in column 2 is illustrated in Figure (4.2), which depicts the health satisfaction difference between "healthy" and "unhealthy" persons for varying values of reference-group health. This graph compares individuals with high (mean + 1 SD) and low values (mean − 1 SD) of the individual health index who are otherwise similar. The graph shows that decreasing (increasing) values of the reference-group health index reduce (raise) the health satisfaction gap between healthy and sick persons. For instance, when reference-group health worsens by one standard deviation the health-satisfaction gap between healthy and sick individuals decreases



Source: Own calculations based on SOEP v29.

Figure 4.2: Health satisfaction gap between healthy and sick individuals

by roughly 11 percent. This provides evidence that better reference-group health aggravates the negative health satisfaction impact of health problems.

The findings of the fixed effects models are presented in Table (4.5). The fixed effects model exploits within-individual variation and thus provides more consistent estimates than the OLS model. The coefficient now measures the influence of transitory changes in reference-group health, or the relative health distance, on individual health satisfaction. Compared to the OLS estimates, the fixed effects coefficients decrease substantially. This can be explained by both unobserved permanent differences across individuals and attenuation bias due to low time-series variation of reference-group health. The association between reference-group health and individual health satisfaction is now insignificant across all specifications. However, we continue to find a significant health norm effect in the second column. The point estimate for the interaction term is quite similar to the OLS result. The corresponding reduction of the health satisfaction gap between healthy and sick individuals is also similar to the OLS finding and amounts to 11 percent.

All in all, our empirical findings indicate quantitatively unimportant effects of relative health on individual health satisfaction. How individuals evaluate their health predominantly rather depends on own health status, and less on how it compares to the health of similar others. In all specifications, the point estimates on individual health status are considerably larger than the coefficients on reference-group health. Admittedly, we obtain a significant interaction effect between own health and reference-group health. However, the corresponding health norm effect appears to be rather small.

4.4.2 Sensitivity Analysis

Refining the Reference Group

This study assumes that respondents make interpersonal comparisons when evaluating their health status. Still, the critical question is who they compare themselves to. Based on survey evidence regarding the most important comparison groups in the subjective assessment of health

Table 4.5: Health satisfaction and relative health, fixed effects results

	All individuals		Healthier	Sicker
	(1)	(2)	(3)	(4)
Own health	0.490***	0.494***	0.524***	0.459***
	(0.005)	(0.005)	(0.011)	(0.008)
Ref.-group health/10	0.015	0.045	-0.022	-0.049
	(0.061)	(0.061)	(0.087)	(0.103)
Own health × Ref.-group health		0.030***		
		(0.004)		
Control variables	Yes	Yes	Yes	Yes
Individual FE	Yes	Yes	Yes	Yes
No. of observations	93,674	93,674	52,033	41,641
No. of individuals	25,788	25,788	19,068	17,116

Note: Robust standard errors, clustered at the individual level, in parentheses. *Health satisfaction* is measured with the quasi-cardinal and approximately standardized health-satisfaction scores. *Own health index* and *reference-group health index* are standardized such that their mean equals zero and their standard deviation equals one. Control variables include sex, age, net household income, years of education, number of adults in the household, number of children in the household, employment status, marital status, a dummy variable for East Germans, and dummy variables for 16 states and 6 survey years. Significance levels: *** p<0.01, ** p<0.05, * p<0.1.

Source: Own calculations based on SOEP v29.

(see Kaplan and Baron-Epel, 2003), the baseline reference-group specification includes persons of the same sex, age, with the same educational background and living in the same region. We nevertheless provide estimates using an alternative, or rather more refined, definition of the reference group. The alternative reference-group specification includes occupation in addition to the baseline attributes. This specification rests on the idea that individuals frequently engage in comparisons with their colleagues to evaluate their situation. Coworkers are often cited as the most important reference group when it comes to income comparisons (e.g. Clark and Senik, 2010). Work-related comparisons may also exist in the domain of health problems (e.g. Buunk et al., 2001). Thus, we follow Pischke (2011) and Blanco-Perez (2012) and set up a reference group which also consists of individuals in the the same occupational group. Occupations are categorized according to the ISCO-88 scheme and aggregated into 22 categories. This leads to a number of 996 reference groups. Average group size amounts to 20 individuals. The minimum number of individuals is equal to 10, the maximum number amounts to 58 individuals. Clearly, this procedure is limited to employed individuals.

Table (4.6) shows the estimates based on the refined reference-group. The number of observations used for the estimation reduces considerably. This can be attributed to the fact that the analysis is restricted to employed individuals, and that there are many cells or reference groups with less than 10 individuals. The implications of these findings are twofold: First, the relative health effect might depend on the choice of the reference group. The coefficient on the interaction term in column 2 is now insignificant at conventional levels. Thus, health norm effects appear to be less relevant in work-related comparisons. Second, the results support the conclusion that reference-group effects are rather unimportant for individual health satisfaction. The size of the point estimates for reference-group health are generally very small and insignificant. However, the results could also be driven by increased attenuation bias. When we use the refined reference-group specification, the number of unique individuals contributing more than two observations to the panel estimation substantially declines. Thus, there is considerably less time-series variation in other people's health than in the baseline sample. This could drive the reference-group effects further down to zero in the fixed effects model.

Table 4.6: Estimation results based on reference groups including occupation

	All individuals		Healthier	Sicker
	(1)	(2)	(3)	(4)
Own health	0.492***	0.492***	0.573***	0.452***
	(0.013)	(0.013)	(0.033)	(0.023)
Ref.-group health/10	-0.090	-0.099	-0.128	0.057
	(0.091)	(0.092)	(0.128)	(0.160)
Own health ×		0.005		
Ref.-group health		(0.009)		
Control variables	Yes	Yes	Yes	Yes
Individual FE	Yes	Yes	Yes	Yes
No. of observations	15,796	15,796	9,210	6,586
No. of individuals	6,747	6,747	4,661	3,719

Note: Robust standard errors, clustered at the individual level, in parentheses. *Health satisfaction* is measured with the quasi-cardinal and approximately standardized health-satisfaction scores. *Own health index* and *reference-group health index* are standardized such that their mean equals zero and their standard deviation equals one. Control variables include sex, age, net household income, years of education, number of adults in the household, number of children in the household, employment status, marital status, a dummy variable for East Germans, and dummy variables for 16 states and 6 survey years. Significance levels: *** $p<0.01$, ** $p<0.05$, * $p<0.1$.

Source: Own calculations based on SOEP v29.

The Role of Physical and Mental Health

This section presents the estimation results when the overall health index is divided into the physical and the mental component. The findings of this exercise are shown in Table (4.7). Panel A illustrates the estimates when only physical health is included, and panel B shows the estimation results based on mental health. The results show that physical health appears to be more important for individual health satisfaction than mental health. We do not have a convincing explanation for this rather surprising finding. One reason could be that the physical component of the SF12 also includes a measure for self-assessed general health status which seems to be highly correlated with health satisfaction. This finding is, however, in line with the estimations by Winkelmann and Studer (forthcoming) who analyze the relationship between physical and mental health on the one hand and health satisfaction on the other hand with SOEP data. Furthermore, the estimations generally support the finding that comparison health is rather unimportant for individual health satisfaction, because the point estimates on the reference-group variables are rather small and in most cases insignificant.

Table 4.7: Estimation results comparing physical and mental health status

	All individuals		Healthier	Sicker
	(1)	(2)	(3)	(4)
	A: Physical health only			
Own health	0.464***	0.468***	0.422***	0.447***
	(0.005)	(0.005)	(0.011)	(0.009)
Ref.-group health/10	0.070	0.139*	0.183	-0.015
	(0.082)	(0.083)	(0.119)	(0.141)
Own health ×		0.046***		
Ref.-group health		(0.004)		
Control variables	Yes	Yes	Yes	Yes
Individual FE	Yes	Yes	Yes	Yes
No. of observations	93,674	93,674	52,220	41,545
No. of individuals	25,788	25,788	19,511	17,502
	B: Mental health only			
Own health	0.166***	0.166***	0.045***	0.201***
	(0.004)	(0.004)	(0.010)	(0.008)
Ref.-group health/10	0.021	0.020	0.067	-0.070
	(0.040)	(0.040)	(0.058)	(0.071)
Own health ×		0.004		
Ref.-group health		(0.003)		
Control variables	Yes	Yes	Yes	Yes
Individual FE	Yes	Yes	Yes	Yes
No. of observations	93,674	93,674	52,204	41,470
No. of individuals	25,788	25,788	19,857	17,917

Note: Robust standard errors, clustered at the individual level, in parentheses. *Health satisfaction* is measured with the quasi-cardinal and approximately standardized health-satisfaction scores. *Own health index* and *reference-group health index* are standardized such that their mean equals zero and their standard deviation equals one. Control variables include sex, age, net household income, years of education, number of adults in the household, number of children in the household, employment status, marital status, a dummy variable for East Germans, and dummy variables for 16 states and 6 survey years. Significance levels: *** $p<0.01$, ** $p<0.05$, * $p<0.1$.

Source: Own calculations based on SOEP v29.

Some differences between physical health and mental health, nevertheless, occur. As illustrated in the second column of Table (4.7), there is a significant interaction between own physical health and reference-group physical health, whereas the interaction term is insignificant for mental health status. The different comparisons effects of physical health and mental health are to some extent consistent with economic research on positional concerns in the context of health. This literature shows that status or norm effects are more widespread in the context of physical health aspects (e.g. Blanchflower et al., 2009; Carrieri and De Paola, 2012; Mujcic and Frijters, 2015).

Panel Attrition

This section presents estimates based on a fully balanced sample of individuals. Health-related panel attrition could influence the estimates on individual health status, the average health in the reference group and thus the effect of peer-group health on health satisfaction. To avoid bias due to health-related losses, we restrict the estimation procedure to those respondents who provide information on all variables in all survey years. The corresponding estimation results are presented in Table (4.8). All in all, the estimation results using the balanced sample resemble the baseline findings using the unbalanced sample. The coefficients on reference-group health remain quantitatively small and insignificant. Moreover, the estimated interaction effect between own health and reference group health is very similar to the baseline finding in Table (4.5).

Table 4.8: Estimation results using a balanced sample

	All individuals		Healthier	Sicker
	(1)	(2)	(3)	(4)
Own health	0.480***	0.484***	0.558***	0.437***
	(0.007)	(0.007)	(0.022)	(0.016)
Ref.-group health/10	0.071	0.087	-0.062	0.071
	(0.086)	(0.086)	(0.166)	(0.218)
Own health × Ref.-group health		0.028***		
		(0.006)		
Control variables	Yes	Yes	Yes	Yes
Individual FE	Yes	Yes	Yes	Yes
No. of observations	39,126	39,126	9,492	5,640
No. of individuals	6,521	6,521	1,582	940

Note: Robust standard errors, clustered at the individual level, in parentheses. *Health satisfaction* is measured with the quasi-cardinal and approximately standardized health-satisfaction scores. *Own health index* and *reference-group health index* are standardized such that their mean equals zero and their standard deviation equals one. Control variables include sex, age, net household income, years of education, number of adults in the household, number of children in the household, employment status, marital status, a dummy variable for East Germans, and dummy variables for 16 states and 6 survey years. Significance levels: *** p<0.01, ** p<0.05, * p<0.1.

Source: Own calculations based on SOEP v29.

Fixed Effects Ordered Logit Estimates

In this section, we test the validity of the POLS approach and compare it with an estimator that takes the ordinal nature of the dependent variable into account. Since a fixed effects ordered probit model is unavailable, we use a recently developed method that allows consistent estimation of the fixed effects ordered logit model. Baetschmann et al. (2015) call it the blow-up-and-cluster (BUC) estimator. Accordingly, the estimation sample is enlarged and each individual is replaced by a number of copies equal to the cut-off points of the ordinal health satisfaction variable. Thus, each individual copy reflects a different cut-off point. A dichotomous variable indicates whether the respondent's original health-satisfaction score is equal to or greater than the respective cut-off point. Following this, a conditional fixed effects logit model can be estimated using the entire sample and the dichotomous indicator as the dependent variable. Since the observations are dependent, the standard errors are clustered at the individual level. Baetschmann et al. (2015) have shown that this procedure leads to a consistent estimator of the fixed effects ordered logit model.

Table (4.9) shows the results using the BUC estimator. The coefficients are not directly comparable to the estimates based on the POLS method. However, the BUC findings are qualitatively similar to the POLS estimates. The coefficients for the reference-group health variables are substantially smaller than the point estimates for own health. What is more, the relationship between reference-group health and individual health satisfaction is statistically insignificant across all specifications. Finally, we find a significant health norm effect, as illustrated by the significant and positive interaction term in the second column.

Table 4.9: Estimation results based on fixed effects ordered logit (BUC) models

	All individuals		Healthier	Sicker
	(1)	(2)	(3)	(4)
Own health	1.544***	1.555***	1.782***	1.467***
	(0.017)	(0.017)	(0.041)	(0.030)
Ref.-group health/10	0.136	0.292	-0.056	-0.130
	(0.217)	(0.216)	(0.319)	(0.365)
Own health ×		0.087***		
Ref.-group health		(0.014)		
Control variables	Yes	Yes	Yes	Yes
Individual FE	Yes	Yes	Yes	Yes
No. of observations	248,378	248,378	89,814	93,994
No. of individuals	18,702	18,702	10,827	9,309

Note: Robust standard errors, clustered at the individual level, in parentheses. *Health satisfaction* is measured with the quasi-cardinal and approximately standardized health-satisfaction scores. *Own health index* and *reference-group health index* are standardized such that their mean equals zero and their standard deviation equals one. Control variables include sex, age, net household income, years of education, number of adults in the household, number of children in the household, employment status, marital status, a dummy variable for East Germans, and dummy variables for 16 states and 6 survey years. Significance levels: *** $p<0.01$, ** $p<0.05$, * $p<0.1$.

Source: Own calculations based on SOEP v29.

4.5 Conclusion

Using panel data from the SOEP study, this study examined the role of other people's health status for individual health satisfaction. Specifically, we analyzed how the average health status in the reference group of the respondent influenced satisfaction with health. To measure individual and peer-group health, we used a comprehensive health measure based on the SF-12 health survey which captures both physical and mental health aspects. Furthermore, we assumed that respondents compare their health status with that of other individuals of the same sex, age, education and who lived in the same region. The empirical analysis based on fixed effects models did not reveal strong effects of relative health. Irrespective of the respondent's relative health position, the association between reference-group health and individual health satisfaction was generally insignificant. We found significant but economically negligible health norm effects. A set of sensitivity analyses — including a more refined reference-group specification, a partition of the health index into a physical and mental health scale, a balanced sample, and a variant of the fixed effects ordered logit estimator — supported the general conclusion that relative health effects are rather unimportant for individual health satisfaction.

For one thing, the weak association between reference group health and individual health satisfaction has practical relevance. The empirical results suggest that individual health satisfaction ratings are not confounded by comparison health effects (see also Powdthavee, 2009). Thus, there should be less concerns about health-related externalities or reference-dependent effects when using subjective evaluations of health and quality of life as outcome variables. For another thing, although not conclusive, our evidence may have significance to policy makers deciding on the design and funding of health care provision. The weak empirical relationship between other people's health and individual health satisfaction found in this study lessens the importance of health-related externalities as an argument for public intervention in the health care sector.

The findings and implications of the present study are however subject to a number of limitations. We impose an exogenous reference group based on demographic and socioeconomic characteristics. This approach, which is standard in the economic social comparison literature, is clearly prone to measurement error and does not allow a causal interpretation of the relative health effect. We do not observe the true comparison standard, and there is no information in the general SOEP questionnaire on whether the respondents actually performed interpersonal comparisons with respect to their health status. Furthermore, there is no data on which reference groups the respondents considered when evaluating their medical condition. What is more, even if we measure the true reference group, the estimates could still be biased. As argued by Manski (1993), endogeneity problems arise due to the following factors: First, the individual and the members of his or her peer-group may mutually influence each other, which makes it difficult to determine whether the reference-group affects individuals or vice versa. Furthermore, the health of other people may simply correlate with individual health evaluation because individuals who are part of the same social group face the same institutional or environmental conditions.

Moreover, the comprehensive health index employed in this study is still self-reported, and residual bias due to reporting heterogeneity may still exist. Although being a more reliable measure than condensed measures of self-assessed general health status, income-related heterogeneity in generic measures like the SF-12 may still remain (e.g. Ziebarth, 2010). Another disadvantage of the employed health index is the fact that the health status of others may not be readily observable in larger social groups. Future research should therefore opt for health measures that are more reliable and better observable.

Generally, the estimates provided in this study cannot be interpreted as strictly causal. They rather reflect associations between peer-group health and individual health satisfaction, controlling for the effect of permanent unobserved differences on health satisfaction, health status and social comparisons. Using more sophisticated methods such as instrumental variables or adjacent reference groups, future research should aim for causal analyses of the relative health effect. Furthermore, it would be useful to have cross-country analyses to study the effect of relative health under different social and cultural contexts, because it is likely that cultural differences may influence the way of relative thinking and social comparison orientation (e.g. Gibbons, 1999). Clearly, a better understanding of the (causal) mechanisms is needed to assess the importance of health comparison effects for individuals and policy making.

Appendix

Table 4.A.1: SOEP SF-12 health scales

Physical health scale (*pcs*)	
Physical functioning (2)	*Scale: 1 (greatly) to 3 (not at all)*
State of health affects ascending stairs	When you ascend stairs, i.e. go up several floors on foot: Does your state of health affect you greatly, slightly or not at all?
State of health affects tiring tasks	And what about having to cope with other tiring everyday tasks, i.e. when one has to lift something heavy or when one requires agility: Does your state of health affect you greatly, slightly or not at all?
Role physical (2)	*Scale: 1 (always) to 5 (never)* Please think about the last four weeks. How often did it occur within this period of time, that due to physical health problems ...
Achieved less due to health last 4 weeks	you achieved less than you wanted to at work or in everyday tasks?
Limited due to health last 4 weeks	you were limited in some form at work or in everyday tasks?
Bodily pain (1)	*Scale: 1 (always) to 5 (never)* Please think about the last four weeks. How often did it occur within this period of time, that due to physical health problems thad you had strong physical pains?
General health (1)	*Scale: 1 (very good) to 5 (bad)* How would you describe your current health?
Mental health scale (*mcs*)	
Vitality (1)	*Scale: 1 (always) to 5 (never)* During the last four weeks, how often did you: feel energetic?
Social functioning (1)	*Scale: 1 (always) to 5 (never)* During the last four weeks, how often did you: feel that due to physical and mental health problems your were limited socially, that is, in contact with friends, acquaintances, or relatives?
Role emotional (2)	*Scale: 1 (always) to 5 (never)* During the last for weeks, how often did you: feel that due to mental health or emotional problems...
Achieved less due to mental health the last 4 weeks	you achieved less than you wanted to at work or in everyday activities?
Less thorough due to health last 4 weeks	you carried out your work or everyday tasks less thoroughly than usual?
Mental health (2)	*Scale: 1 (always) to 5 (never)* During the last four weeks, how often did you:...
Run-down, melancholy last 4 weeks	feel down and gloomy?
Well-balanced last 4 weeks	feel calm and relaxed

Source: SOEP v29.

Bibliography

Abadie, A. and Imbens, G. W. (2008). On the failure of the bootstrap for matching estimators, *Econometrica* **76**(6): 1537–1557.

Abel, T. (2008). Cultural capital and social inequality in health, *Journal of Epidemiology and Community Health* **62**(7): e13–e13.

Adler, N. E. and Newman, K. (2002). Socioeconomic disparities in health: Pathways and policies, *Health Affairs* **21**(2): 60–76.

Ai, C. and Norton, E. C. (2003). Interaction terms in logit and probit models, *Economics Letters* **80**(1): 123–129.

Akerlof, G. A. (1980). A theory of social custom, of which unemployment may be one consequence, *Quarterly Journal of Economics* **94**(4): 49–775.

Almlund, M., Duckworth, A. L., Heckman, J. and Kautz, T. (2011). Personality psychology and economics, *in* S. M. Eric A. Hanushek and L. Woessmann (eds), *Handbook of the economics of education*, Vol. 4, North-Holland, Amsterdam, pp. 1–181.

Alpizar, F., Carlsson, F. and Johansson-Stenman, O. (2005). How much do we care about absolute versus relative income and consumption?, *Journal of Economic Behavior & Organization* **56**(3): 405–421.

Altonji, J. G., Elder, T. E. and taber, C. R. (2005). Selection on observed and unobserved variables: Assessing the effectiveness of catholic schools, *Journal of Political Economy* **113**(1): 151–184.

Andersen, H. H., Mühlbacher, A., Nübling, M., Schupp, J. and Wagner, G. G. (2007). Computation of standard values for physical and mental health scale scores using the SOEP version of SF-12v2, *Schmollers Jahrbuch: Journal of Applied Social Science Studies* **127**(1): 171–182.

Andreoni, J. (1990). Impure altruism and donations to public goods: A theory of warm-glow giving, *Economic Journal* **100**(401): 464–477.

Angrist, J. D. and Pischke, J.-S. (2008). *Mostly harmless econometrics: An empiricist's companion*, Princeton University Press, Princeton.

Auld, M. C. and Sidhu, N. (2005). Schooling, cognitive ability and health, *Health Economics* **14**(10): 1019–1034.

Authoring Group Educational Reporting (ed.) (2012). *Bildung in Deutschland 2012 / Education in Germany 2012*, W. Bertelsmann, Bielefeld.

Baetschmann, G., Staub, K. E. and Winkelmann, R. (2015). Consistent estimation of the fixed effects ordered logit model, *Journal of the Royal Statistical Society: Series A (Statistics in Society)* **178**(3): 685–703.

Bang, H. and Robins, J. M. (2005). Doubly robust estimation in missing data and causal inference models, *Biometrics* **61**(4): 962–973.

Barrow, S. and Harrison, R. A. (2005). Unsung heroes who put their lives at risk? Informal caring, health and neighbourhood attachment, *Journal of Public Health* **27**(3): 292–297.

Bauer, J. M. and Sousa-Poza, A. (2015). Impacts of informal caregiving on caregiver employment, health, and family, *Journal of Population Ageing* **8**(3): 1–33.

Bauernschuster, S., Falck, O. and Woessmann, L. (2014). Surfing alone? The internet and social capital: Evidence from an unforeseeable technological mistake, *Journal of Public Economics* **117**: 73–89.

Berkman, L. F., Glass, T., Brissette, I. and Seeman, T. E. (2000). From social integration to health: Durkheim in the new millennium, *Social Science & Medicine* **51**(6): 843–857.

Black, D. A. and Smith, J. A. (2004). How robust is the evidence on the effects of college quality? Evidence from matching, *Journal of Econometrics* **121**(1): 99–124.

Blanchflower, D. G., Landeghem, B. and Oswald, A. J. (2009). Imitative obesity and relative utility, *Journal of the European Economic Association* **7**(2-3): 528–538.

Blanco-Perez, C. (2012). Rethinking the relative income hypothesis, *SOEPpaper No. 501*, DIW Berlin.

Bobinac, A., van Exel, N. J. A., Rutten, F. F. and Brouwer, W. B. (2010). Caring for and caring about: Disentangling the caregiver effect and the family effect, *Journal of Health Economics* **29**(4): 549–556.

Bolin, K., Lindgren, B., Lindström, M. and Nystedt, P. (2003). Investments in social capital—implications of social interactions for the production of health, *Social Science & Medicine* **56**(12): 2379–2390.

Borgonovi, F. (2004). Performing arts attendance: An economic approach, *Applied Economics* **36**(17): 1871–1885.

Bourdieu, P. (1984). *Distinction: A social critique of the judgement of taste*, Harvard University Press, Cambridge, MA.

Burr, J. A., Choi, N. G., Mutchler, J. E. and Caro, F. G. (2005). Caregiving and volunteering: Are private and public helping behaviors linked?, *Journals of Gerontology Series B: Psychological Sciences and Social Sciences* **60**(5): S247–S256.

Buunk, B. and Gibbons, F. X. (1997). Social comparisons in health and illness: A historical overview, *in* B. Buunk and F. X. Gibbons (eds), *Health, coping and well-being: Perspectives from social comparison theory*, Lawrence Erlbaum Associates, Mahwah, NJ, pp. 1–23.

Buunk, B. P., Collins, R. L., Taylor, S. E., VanYperen, N. W. and Dakof, G. A. (1990). The affective consequences of social comparison: Either direction has its ups and downs, *Journal of Personality and Social Psychology* **59**(6): 1238–1249.

Buunk, B. P., Ybema, J. F., Gibbons, F. X. and Ipenburg, M. (2001). The affective consequences of social comparison as related to professional burnout and social comparison orientation, *European Journal of Social Psychology* **31**(4): 337–351.

Bygren, L. O., Weissglas, G., Wikström, B.-M., Konlaan, B. B., Grjibovski, A., Karlsson, A.-B., Andersson, S.-O. and Sjöström, M. (2009). Cultural participation and health: A randomized controlled trial among medical care staff, *Psychosomatic Medicine* **71**(4): 469–473.

Caliendo, M. and Kopeinig, S. (2008). Some practical guidance for the implementation of propensity score matching, *Journal of Economic Surveys* **22**(1): 31–72.

Cannuscio, C. C., Colditz, G. A., Rimm, E. B., Berkman, L. F., Jones, C. P. and Kawachi, I. (2004). Employment status, social ties, and caregivers' mental health, *Social Science & Medicine* **58**(7): 1247–1256.

Carlsson, F., Johansson-Stenman, O. and Martinsson, P. (2007). Do you enjoy having more than others? Survey evidence of positional goods, *Economica* **74**(296): 586–598.

Carpiano, R. M. (2008). Actual or potential neighborhood resources and access to them: Testing hypotheses of social capital for the health of female caregivers, *Social Science & Medicine* **67**(4): 568–582.

Carrieri, V. (2012). Social comparison and subjective well-being: Does the health of others matter?, *Bulletin of Economic Research* **64**(1): 31–55.

Carrieri, V. and De Paola, M. (2012). Height and subjective well-being in Italy, *Economics & Human Biology* **10**(3): 289–298.

Christakis, N. A. and Fowler, J. H. (2007). The spread of obesity in a large social network over 32 years, *New England Journal of Medicine* **357**(4): 370–379.

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

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

Clark, A. and Lelkes, O. (2006). Deliver us from evil: Religion as insurance, *Papers on Economics of Religion PER 06/03*, Department of Economic Theory and Economic History of the University of Granada.

Cohen, S. and Wills, T. (1985). Stress, social support, and the buffering hypothesis., *Psychological Bulletin* **98**(2): 310–357.

Coleman, J. S. (1988). Social capital in the creation of human capital, *American Journal of Sociology* **94**(suppl.): S95–S120.

Colombo, F., Llena-Nozal, A., Mercier, J. and Tjadens, F. (2011). Help wanted? Providing and paying for long-term care, *OECD Health Policy Studies*, OECD Publishing, Paris.

Costa-Font, J. and Mladovsky, P. (2008). Social capital and the social formation of health-related preferences and behaviours, *Health Economics, Policy and Law* **3**(04): 413–427.

Costa, P. T. and McCrae, R. R. (1988). Personality in adulthood: A six-year longitudinal study of self-reports and spouse ratings on the neo personality inventory., *Journal of Personality and Social Psychology* **54**(5): 853–863.

Culyer, A. J. (1971). The nature of the commodity 'health care' and its efficient allocation, *Oxford Economic Papers* **23**(2): 189–211.

Culyer, A. J. and Simpson, H. (1980). Externality models and health: A Rückblick over the last twenty years, *Economic Record* **56**(154): 222–230.

Cutler, D. M., Lleras-Muney, A. and Vogl, T. (2011). Socioeconomic status and health: Dimensions and mechanisms, *in* S. Glied and P. C. Smith (eds), *The Oxford handbook of health economics*, Oxford University Press, Oxford, pp. 124–136.

Cutler, D. M. and Richardson, E. (1997). Measuring the health of the U.S. population, *Brookings Papers on Economic Activity. Microeconomics* **28**: 217–282.

Cuypers, K., Krokstad, S., Holmen, T. L., Knudtsen, M. S., Bygren, L. O. and Holmen, J. (2012). Patterns of receptive and creative cultural activities and their association with perceived health, anxiety, depression and satisfaction with life among adults: The HUNT study, Norway, *Journal of Epidemiology and Community Health* **66**(8): 698–703.

Deaton, A. (2003). Health, inequality, and economic development, *Journal of Economic Literature* **41**(1): 113–158.

Dehejia, R., DeLeire, T. and Luttmer, E. F. (2007). Insuring consumption and happiness through religious organizations, *Journal of Public Economics* **91**(1–2): 259–279.

Dehejia, R. H. and Wahba, S. (1999). Causal effects in nonexperimental studies: Reevaluating the evaluation of training programs, *Journal of the American Statistical Association* **94**(448): 1053–1062.

Deri, C. (2005). Social networks and health service utilization, *Journal of Health Economics* **24**(6): 1076–1107.

DiMaggio, P. (1982). Cultural capital and school success: The impact of status culture participation on the grades of U.S. high school students, *American Sociological Review* **47**(2): 189–201.

Duesenberry, J. S. (1949). *Income, saving, and the theory of consumer behavior*, Harvard University Press, Cambridge.

Durlauf, S. N. and Fafchamps, M. (2005). Social capital, *in* P. Aghion and S. N. Durlauf (eds), *Handbook of Economic Growth*, Vol. 1, Part B, North-Holland, Amsterdam, pp. 1639–1699.

Durlauf, S. N. and Ioannides, Y. M. (2010). Social interactions, *Annual Review of Economics* **2**(1): 451–478.

Easterlin, R. A. (1974). Does economic growth improve the human lot? Some empirical evidence, *in* P. A. David and M. W. Reder (eds), *Nations and households in economic growth. Essays in honor of Moses Abramovitz*, Academic Press, New York, pp. 89–125.

Eckersley, R. (2015). Beyond inequality: Acknowledging the complexity of social determinants of health, *Social Science & Medicine* **147**: 121–125.

Eibich, P. and Ziebarth, N. R. (2014). Examining the structure of spatial health effects in Germany using hierarchical Bayes models, *Regional Science and Urban Economics* **49**: 305–320.

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

Elster, J. (1989). Social norms and economic theory, *Journal of Economic Perspectives* **3**(4): 99–117.

Elster, J. (1991). *Envy in social life*, MIT Press, Cambridge.

Falk, A. and Knell, M. (2004). Choosing the Joneses: Endogenous goals and reference standards, *Scandinavian Journal of Economics* **106**(3): 417–435.

Fehr, E. and Schmidt, K. M. (1999). A theory of fairness, competition, and cooperation, *Quarterly Journal of Economics* **114**(3): 817–868.

Ferrer-i-Carbonell, A. (2005). Income and well-being: An empirical analysis of the comparison income effect, *Journal of Public Economics* **89**(5): 997–1019.

Festinger, L. (1954). A theory of social comparison processes, *Human Relations* **7**(2): 117–140.

Fienberg, S. E., Loftus, E. F. and Tanur, J. M. (1985). Cognitive aspects of health survey methodology: An overview, *Milbank Memorial Fund Quarterly. Health and Society* **63**(3): 547–564.

Folland, L. and Rocco, E. (eds) (2014). *The economics of social capital and health. A conceptual and empirical roadmap*, World Scientific Publishing, Singapore.

Folland, S. (2008). An economic model of social capital and health, *Health Economics, Policy and Law* **3**(04): 333–348.

Fowler, J. H. and Christakis, N. A. (2008). Dynamic spread of happiness in a large social network: Longitudinal analysis over 20 years in the Framingham Heart Study, *British Medical Journal* **337**(7685): 23–27.

Frank, R. H. (1985). The demand for unobservable and other nonpositional goods, *American Economic Review* **75**(1): 101–116.

Frank, R. H. (2008). Should public policy respond to positional externalities?, *Journal of Public Economics* **92**(8): 1777–1786.

Frey, B. S. (2008). Cities, culture and happiness: The arts as a means to enhance indivdiual well-being can have an impact on how culture should be encouraged by policy, *ETHOS* **June**(World Cities Summit Issue): 102–111.

Frick, J. R. and Ziebarth, N. R. (2013). Welfare-related health inequality: Does the choice of measure matter?, *European Journal of Health Economics* **14**(3): 431–442.

Frijters, P., Haisken-DeNew, J. P. and Shields, M. A. (2011). The increasingly mixed proportional hazard model: An application to socioeconomic status, health shocks, and mortality, *Journal of Business & Economic Statistics* **29**(2): 271–281.

Ganzeboom, H. B. G. (1984). Explaining differential participation in high-cultural activities — A confrontation of information-processing and status-seeking theories, *in* W. Raub (ed.), *Theoretical models and empirical analyses. Contributions to the explanation of individual actions and collective phenomena*, E.S. Publications, Utrecht, pp. 186–203.

Gelman, A. (2008). Scaling regression inputs by dividing by two standard deviations, *Statistics in Medicine* **27**(15): 2865–2873.

Gelman, A. and Hill, J. (2006). *Data analysis using regression and multilevel/hierarchical models*, Cambridge University Press, Cambridge.

German Federal Parliament (2007). Schlussbericht der Enquete-Kommission "Kultur in Deutschland" / Final report of the enquiry on "Culture in Germany", *Printed matter 16/7000*, German Federal Parliament, 16th legislative period.

Gibbons, F. (1999). Social comparison as a mediator of response shift, *Social Science & Medicine* **48**(11): 1517–1530.

Gibbons, F. X. and Buunk, B. P. (1999). Individual differences in social comparison: Development of a scale of social comparison orientation., *Journal of Personality and Social Psychology* **76**(1): 129–142.

Glaeser, E. L., Laibson, D. and Sacerdote, B. (2002). An economic approach to social capital, *Economic Journal* **112**(483): F437–F458.

Graham, C. (2009). *Happiness around the world: The paradox of happy peasants and miserable millionaires*, Oxford University Press, Oxford.

Gravelle, H. and Sutton, M. (2009). Income, relative income, and self-reported health in Britain 1979–2000, *Health Economics* **18**(2): 125–145.

Gray, C. M. (1998). Hope for the future? Early exposure to the arts and adult visits to art museums, *Journal of Cultural Economics* **22**(2-3): 87–98.

Gray, C. M. (2011). Participation, *in* R. Towse (ed.), *A Handbook in Cultural Economics*, 2nd edn, Edward Elgar, Cheltenham, pp. 326–333.

Greene, W. H. and Hensher, D. A. (2010). *Modeling ordered choices: A primer*, Cambridge University Press, Cambridge.

Groot, W. (2000). Adaptation and scale of reference bias in self-assessments of quality of life, *Journal of Health Economics* **19**(3): 403–420.

Groot, W. (2003). Scale of reference bias and the evolution of health, *European Journal of Health Economics* **4**(3): 176–183.

Groot, W. and Van Den Brink, H. M. (2003). Sympathy and the value of health: The spill-over effects of migraine on household well-being, *Social Indicators Research* **61**(1): 97–120.

Grossman, M. (1972). On the concept of health capital and the demand for health, *Journal of Political Economy* **80**(2): 223–255.

Hamermesh, D. S. and Biddle, J. E. (1994). Beauty and the labor market, *American Economic Review* **84**(5): 1174–94.

Hank, K. and Stuck, S. (2008). Volunteer work, informal help, and care among the 50+ in Europe: Further evidence for linked productive activities at older ages, *Social Science Research* **37**(4): 1280–1291.

Hansen, T. and Slagsvold, B. (2015). Feeling the squeeze? The effects of combining work and informal caregiving on psychological well-being, *European Journal of Ageing* **12**(1): 51–60.

Hansen, T., Slagsvold, B. and Ingebretsen, R. (2013). The strains and gains of caregiving: An examination of the effects of providing personal care to a parent on a range of indicators of psychological well-being, *Social Indicators Research* **114**(2): 323–343.

Heckman, J. J., Ichimura, H. and Todd, P. (1998). Matching as an econometric evaluation estimator, *Review of Economic Studies* **65**(2): 261–294.

Helliwell, J. F. and Putnam, R. D. (2004). The social context of well-being, *Philosophical Transactions of the Royal Society of London B: Biological Sciences* **359**(1449): 1435–1446.

Hertzog, C., Kramer, A. F., Wilson, R. S. and Lindenberger, U. (2008). Enrichment effects on adult cognitive development. Can the functional capacity of older adults be preserved and enhanced?, *Psychological Science in the Public Interest* **9**(1): 1–65.

Hille, A. and Schupp, J. (2015). How learning a musical instrument affects the development of skills, *Economics of Education Review* **44**: 56–82.

Hillesheim, I. and Mechtel, M. (2013). How much do others matter? Explaining positional concerns for different goods and personal characteristics, *Journal of Economic Psychology* **34**: 61–77.

Ho, D. E., Imai, K., King, G. and Stuart, E. A. (2007). Matching as nonparametric preprocessing for reducing model dependence in parametric causal inference, *Political Analysis* **15**(3): 199–236.

Huang, J., Van den Brink, H. M. and Groot, W. (2009). A meta-analysis of the effect of education on social capital, *Economics of Education Review* **28**(4): 454–464.

Hurley, J. and Mentzakis, E. (2013). Health-related externalities: Evidence from a choice experiment, *Journal of Health Economics* **32**(4): 671–681.

Hyyppä, M. T. (2010). How does cultural participation contribute to social capital and well-being?, *in* M. T. Hyyppä (ed.), *Healthy Ties. Social capital, population health and survival*, Springer, Dordrecht, pp. 43–53.

Ichino, A., Mealli, F. and Nannicini, T. (2008). From temporary help jobs to permanent employment: What can we learn from matching estimators and their sensitivity?, *Journal of Applied Econometrics* **23**(3): 305–327.

Imbens, G. W. (2004). Nonparametric estimation of average treatment effects under exogeneity: A review, *Review of Economics and Statistics* **86**(1): 4–29.

Imbens, G. W. (2015). Matching methods in practice: Three examples, *Journal of Human Resources* **50**(2): 373–419.

Imbens, G. W. and Wooldridge, J. M. (2009). Recent developments in the econometrics of program evaluation, *Journal of Economic Literature* **47**(1): 5–86.

Iwasaki, Y., Mannell, R. C., Smale, B. J. and Butcher, J. (2005). Contributions of leisure participation in predicting stress coping and health among police and emergency response services workers, *Journal of Health Psychology* **10**(1): 79–99.

Jacobs, K., Kuhlmey, A., Greß, S., Klauber, S. and Schwinger, A. (eds) (2016). *Pflege-Report 2016. Schwerpunkt: Die Pflegenden im Fokus / Care report 2016. Focus on informal caregivers*, Schattauer, Stuttgart.

Jeannotte, M. S. (2003). Singing alone? The contribution of cultural capital to social cohesion and sustainable communities, *International Journal of Cultural Policy* **9**(1): 35–49.

Johansson, S. E., Konlaan, B. B. and Bygren, L. O. (2001). Sustaining habits of attending cultural events and maintenance of health: A longitudinal study, *Health Promotion International* **16**(3): 229–234.

Jones, A. M. and Wildman, J. (2008). Health, income and relative deprivation: Evidence from the BHPS, *Journal of Health Economics* **27**(2): 308–324.

Jürges, H. (1998). Beruflich bedingte Umzüge von Doppelverdienern. Eine empirische Analyse mit Daten des SOEP / Vocationally motivated migration behavior in double-income households. An empirical analysis using SOEP data, *Zeitschrift für Soziologie* **27**(5): 358–377.

Kaplan, G. and Baron-Epel, O. (2003). What lies behind the subjective evaluation of health status?, *Social Science & Medicine* **56**(8): 1669–1676.

Kawachi, I. and Berkman, L. F. (2001). Social ties and mental health, *Journal of Urban Health* **78**(3): 458–467.

Kawachi, I., Subramanian, S. and Kim, D. (2008). Social capital and health, *in* I. Kawachi, S. Subramanian and D. Kim (eds), *Social Capital and health*, Springer, New York, pp. 1–26.

Khawaja, M. and Mowafi, M. (2006). Cultural capital and self-rated health in low income women: Evidence from the Urban Health Study, Beirut, Lebanon, *Journal of Urban Health* **83**(3): 444–458.

Kim, S. and Kim, H. (2009). Does cultural capital matter? Cultural divide and quality of life, *Social Indicators Research* **93**(2): 295–313.

Kling, J. R., Liebman, J. B. and Katz, L. F. (2007). Experimental analysis of neighborhood effects, *Econometrica* **75**(1): 83–119.

Konlaan, B. B., Björby, N., Bygren, L. O., Weissglas, G., Karlsson, L. G. and Widmark, M. (2000). Attendance at cultural events and physical exercise and health: A randomized controlled study, *Public Health* **114**(5): 316–319.

Kraaykamp, G. and Eijck, K. v. (2005). Personality, media preferences, and cultural participation, *Personality and Individual Differences* **38**(7): 1675–1688.

Labelle, R. J. and Hurley, J. E. (1992). Implications of basing health-care resource allocations on cost-utility analysis in the presence of externalities, *Journal of Health Economics* **11**(3): 259–277.

Lechner, M. (2008). A note on endogenous control variables in causal studies, *Statistics & Probability Letters* **78**(2): 190–195.

Lechner, M. (2009). Long-run labour market and health effects of individual sports activities, *Journal of Health Economics* **28**(4): 839–854.

Lechner, M. and Sari, N. (2015). Labor market effects of sports and exercise: Evidence from Canadian panel data, *Labour Economics* **35**: 1–15.

Leuven, E. and Sianesi, B. (2003). PSMATCH2: Stata module to perform full Mahalanobis and propensity score matching, common support graphing, and covariate imbalance testing. This version: 4.0.11.
URL:`http://ideas.repec.org/c/boc/bocode/s432001.html`

Lévy-Garboua, L. and Montmarquette, C. (1996). A microeconometric study of theatre demand, *Journal of Cultural Economics* **20**(1): 25–50.

Lindeboom, M. (2006). Health and work of older workers, *in* A. M. Jones (ed.), *The Elgar companion to health economics*, Edward Elgar Publishing, Cheltenham, pp. 24–35.

Mackenbach, J. P., Stirbu, I., Roskam, A.-J. R., Schaap, M. M., Menvielle, G., Leinsalu, M. and Kunst, A. E. (2008). Socioeconomic inequalities in health in 22 European countries, *New England Journal of Medicine* **358**(23): 2468–2481.

MacKinnon, J. G. (2006). Bootstrap methods in econometrics, *Economic Record* **82**(s1): S2–S18.

Mallick, D. (2009). Marginal and interaction effects in ordered response models, *MPRA Paper No. 9617*, University of Munich.

Manski, C. F. (1993). Identification of endogenous social effects: The reflection problem, *Review of Economic Studies* **60**(3): 531–542.

Manski, C. F. (2000). Economic analysis of social interactions, *Journal of Economic Perspectives* **14**(3): 115–136.

Marcus, J. (2014). Does job loss make you smoke and gain weight?, *Economica* **81**(324): 626–648.

McCarthy, K. F., Ondaatje, E. H., Zakaras, L. and Brooks, A. (2004). *Gifts of the muse: Reframing the debate about the benefits of the arts*, RAND Corporation, Santa Monica.

Morgan, S. L. and Harding, D. J. (2006). Matching estimators of causal effects: Prospects and pitfalls in theory and practice, *Sociological Methods & Research* **35**(1): 3–60.

Morgan, S. L. and Winship, C. (2007). *Counterfactuals and causal inference: Methods and principles for social research*, Cambridge University Press, Cambridge.

Morrison, W. G. and West, E. G. (1986). Child exposure to the performing arts: The implications for adult demand, *Journal of Cultural Economics* **10**(1): 17–24.

Mujcic, R. and Frijters, P. (2015). Conspicuous consumption, conspicuous health, and optimal taxation, *Journal of Economic Behavior & Organization* **111**: 59–70.

Nannicini, T. et al. (2007). Simulation-based sensitivity analysis for matching estimators, *Stata Journal* **7**(3): 334–350.

Notten, N., Lancee, B., van de Werfhorst, H. and Ganzeboom, H. (2015). Educational stratification in cultural participation: Cognitive competence or status motivation?, *Journal of Cultural Economics* **39**(2): 177–203.

Oswald, A. J. and Powdthavee, N. (2007). Obesity, unhappiness, and the challenge of affluence: Theory and evidence, *Economic Journal* **117**(6): 441–454.

Paolucci, F. (2011). *Health care financing and insurance: Options for design*, Springer, Berlin, chapter Economic rationales for the design of health care financing schemes, pp. 13–32.

Pearlin, L. I., Mullan, J. T., Semple, S. J. and Skaff, M. M. (1990). Caregiving and the stress process: An overview of concepts and their measures, *The Gerontologist* **30**(5): 583–594.

Pfaff, S. (2012). Pendeln oder umziehen? Mobilitätsentscheidungen in Deutschland zwischen 2000 und 2009 / Commuting or relocation? Mobility decisions in Germany between 2000 and 2009, *Zeitschrift für Soziologie* **41**(6): 458–477.

Pinquart, M. and Sörensen, S. (2003). Differences between caregivers and noncaregivers in psychological health and physical health: A meta-analysis., *Psychology and Aging* **18**(2): 250–267.

Pinquart, M. and Sörensen, S. (2006). Gender differences in caregiver stressors, social resources, and health: An updated meta-analysis, *Journals of Gerontology Series B: Psychological Sciences and Social Sciences* **61**(1): P33–P45.

Pischke, J.-S. (2011). Money and happiness: Evidence from the industry wage structure, *NBER Working Paper No. 17056*, National Bureau of Economic Research.

Powdthavee, N. (2009). Ill-health as a household norm: Evidence from other people's health problems, *Social Science & Medicine* **68**(2): 251–259.

Putnam, R. D. (2001). *Bowling alone: The collapse and revival of American community*, Simon and Schuster, New York.

Renton, A., Phillips, G., Daykin, N., Yu, G., Taylor, K. and Petticrew, M. (2012). Think of your art-eries: Arts participation, behavioural cardiovascular risk factors and mental well-being in deprived communities in London, *Public Health* **126**: S57–S64.

Richter, D., Metzing, M., Weinhardt, M. and Schupp, J. (2013). SOEP scales manual, *SOEP Survey Papers 138: Series C*, DIW/SOEP, Berlin.

Robert Koch Institute (2015). Informal caregivers — Germany's largest nursing service, *GBE Kompakt* **6**(3): 1–10.

Rocco, L. and Fumagalli, E. (2014). The empirics of social capital and health, *in* S. Folland and L. Rocco (eds), *The economics of social capital and health. A conceptual and empirical roadmap*, World Scientific Publishing, Singapore, pp. 93–152.

Rodakowski, J., Skidmore, E. R., Rogers, J. C. and Schulz, R. (2012). Role of social support in predicting caregiver burden, *Archives of Physical Medicine and Rehabilitation* **93**(12): 2229–2236.

Rosenbaum, P. R. (1984). The consquences of adjustment for a concomitant variable that has been affected by the treatment, *Journal of the Royal Statistical Society. Series A (General)* **147**(5): 656–666.

Rosenbaum, P. R. and Rubin, D. B. (1983). The central role of the propensity score in observational studies for causal effects, *Biometrika* **70**(1): 41–55.

Rosenbaum, P. R. and Rubin, D. B. (1985). Constructing a control group using multivariate matched sampling methods that incorporate the propensity score, *American Statistician* **39**(1): 33–38.

Ross, C. E. and Wu, C.-L. (1995). The links between education and health, *American Sociological Review* **60**(5): 719–745.

Rubin, R. M. and White-Means, S. I. (2009). Informal caregiving: Dilemmas of sandwiched caregivers, *Journal of Family and Economic Issues* **30**(3): 252–267.

Runciman, W. G. (1966). *Relative deprivation and social justice: A study of attitudes to social inequality in twentieth-century England*, University of California Press, Berkeley.

Schmitz, H. and Stroka, M. A. (2013). Health and the double burden of full-time work and informal care provision — Evidence from administrative data, *Labour Economics* **24**: 305–322.

Schmitz, H. and Westphal, M. (2015). Short- and medium-term effects of informal care provision on female caregivers' health, *Journal of Health Economics* **42**: 174–185.

Schulz, R. and Sherwood, P. R. (2008). Physical and mental health effects of family caregiving, *American Journal of Nursing* **108**(9 Suppl): 23–27.

Seaman, B. (2006). Empirical studies of demand for the performing arts, *in* V. A. Ginsburg and D. Throsby (eds), *Handbook of the economics of art and culture*, Vol. 1, North-Holland, Amsterdam, pp. 415–472.

Smith, J. A. and Todd, P. E. (2005). Does matching overcome LaLonde's critique of nonexperimental estimators?, *Journal of Econometrics* **125**(1): 305–353.

Solnick, S. J. and Hemenway, D. (1998). Is more always better? A survey on positional concerns, *Journal of Economic Behavior & Organization* **37**(3): 373–383.

Solnick, S. J. and Hemenway, D. (2005). Are positional concerns stronger in some domains than in others?, *American Economic Review* **95**(2): 147–151.

Statistical Offices of the Federation and the Länder (ed.) (2012). *Kulturfinanzbericht 2012 / Cultural finance report 2012*, Federal Statistical Office, Wiesbaden.

Stigler, G. J. and Becker, G. S. (1977). De gustibus non est disputandum, *American Economic Review* **67**(2): 76–90.

Stine-Morrow, E. A., Parisi, J. M., Morrow, D. G., Greene, J. and Park, D. C. (2007). An engagement model of cognitive optimization through adulthood, *Journals of Gerontology Series B: Psychological Sciences and Social Sciences* **62**(Special Issue 1): 62–69.

Stouffer, S. A., Lumsdaine, A. A., Lumsdaine, M. H., Williams Jr. , Robin M., Smith, M. Brewster, Janis, I. L., Star, S. A. and Cottrell Jr. , Leonard S. (1949). *The American soldier: combat and its aftermath. (Studies in social psychology in World War II, Vol. 2.)*, Princeton University Press, Princeton.

Stronks, K., van de Mheen, H., Looman, C. W. and Mackenbach, J. P. (1998). The importance of psychosocial stressors for socio-economic inequalities in perceived health, *Social Science & Medicine* **46**(4): 611–623.

Thoits, P. A. (2011). Mechanisms linking social ties and support to physical and mental health, *Journal of Health and Social Behavior* **52**(2): 145–161.

Throsby, D. (1994). The production and consumption of the arts: A view of cultural economics, *Journal of Economic Literature* **32**(1): 1–29.

Umberson, D. and Montez, J. K. (2010). Social relationships and health: A flashpoint for health policy, *Journal of Health and Social behavior* **51**(1 suppl): S54–S66.

Van den Berg, B. and Ferrer-i Carbonell, A. (2007). Monetary valuation of informal care: The well-being valuation method, *Health Economics* **16**(11): 1227–1244.

Van Oorschot, W. and Arts, W. (2005). The social capital of European welfare states: The crowding out hypothesis revisited, *Journal of European Social Policy* **15**(1): 5–26.

Van Praag, B. M. and Ferrer-i-Carbonell, A. (2008). *Happiness quantified: a satisfaction calculus approach*, Oxford University Press, Oxford.

Veblen, T. (1899). *The theory of the leisure class: An economic study of institutions*, Macmillan, London.

Viscusi, W. K., Magat, W. A. and Forrest, A. (1988). Altruistic and private valuations of risk reduction, *Journal of Policy Analysis and Management* **7**(2): 227–245.

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

Wagstaff, A. and Van Doorslaer, E. (2000). Income inequality and health: What does the literature tell us?, *Annual Review of Public Health* **21**(1): 543–567.

Wilkinson, A. V., Waters, A. J., Bygren, L. O. and Tarlov, A. R. (2007). Are variations in rates of attending cultural activities associated with population health in the United States?, *BMC Public Health* **7**(1): 226.

Wilkinson, R. G. (1999). Health, hierarchy, and social anxiety, *Annals of the New York Academy of Sciences* **896**(1): 48–63.

Wilkinson, R. G. and Pickett, K. E. (2006). Income inequality and population health: A review and explanation of the evidence, *Social Science & Medicine* **62**(7): 1768–1784.

Wills, T. A. (1981). Downward comparison principles in social psychology, *Psychological Bulletin* **90**(2): 245–271.

Wilson, S. E. (2002). The health capital of families: An investigation of the inter-spousal correlation in health status, *Social Science & Medicine* **55**(7): 1157–1172.

Winkelmann, R. (2009). Unemployment, social capital, and subjective well-being, *Journal of Happiness Studies* **10**(4): 421–430.

Winkelmann, R. and Studer, R. (forthcoming). Econometric analysis of ratings — with an application to health and wellbeing, *Swiss Journal of Economics and Statistics* .

Wunder, C. (2009). Income inequality and job satisfaction of full-time employees in Germany, *Journal of Income Distribution* **18**(2): 70–91.

Yitzhaki, S. (1979). Relative deprivation and the Gini coefficient, *Quarterly Journal of Economics* **93**(2): 321–324.

Ziebarth, N. (2010). Measurement of health, health inequality, and reporting heterogeneity, *Social Science & Medicine* **71**(1): 116–124.

Zuchandke, A., Reddemann, S. and Krummaker, S. (2011). Financing long-term care in Germany, *in* J. Costa-Font and C. Courbage (eds), *Financing long-term care in Europe: Institutions, markets and models*, Palgrave Macmillan, London, pp. 214–235.