767
2015

SOEPpapers
on Multidisciplinary Panel Data Research

SOEP — The German Socio-Economic Panel study at DIW Berlin

767-2015

Leave the Drama on the Stage: The Effect of Cultural Participation on Health

Lars Thiel

ISSN: 1864-6689 (online)

German Socio-Economic Panel Study (SOEP)
DIW Berlin
Mohrenstrasse 58
10117 Berlin, Germany

Contact: Uta Rahmann | soeppapers@diw.de

Leave the Drama on the Stage: The Effect of Cultural Participation on Health

Lars Thiel*

Schumpeter School of Business and Economics, University of Wuppertal

19th June 2015

Abstract

The aim of this study is to estimate the causal effect of cultural participation on health status. Cultural activities may directly impact upon health through palliative coping or substituting health-compromising behaviors. Cultural engagement may also facilitate the development of social networks, which can improve health via social support and the dissemination of social health norms. Previous estimates on the arts-health relationship are potentially biased due to reverse causality and unobserved heterogeneity. Using individual-level data from Germany, we employ propensity-score matching methods. The treatment group is confined to individuals that visit cultural events at least once a month. The participation equation includes a rich set of personal characteristics that cover the respondents' demographic and social background, social capital and leisure-time activities, health-related lifestyle, personality and childhood environment. We explicitly consider reverse causality by including the pre-treatment trends in health outcomes among the covariates. To deal with time-fixed unobserved heterogeneity, we combine the matching model with a difference-in-difference approach. We find that frequent cultural-event visits are unrelated to health once we account for unobserved persistent differences across individuals. However, examining the dose-response relationship we find positive mental-health effects of low levels of cultural participation compared to non-attendance. Our results may thus yield important insights on the effectiveness of arts participation as a means to reduce social inequalities in health.

Keywords: Cultural participation; mental health; physical health; propensity-score matching; multivalued treatment

JEL Classification: I12, Z11

*Address: Rainer-Gruenter-Str. 21, FN.01.03, 42119 Wuppertal, Germany, telephone: +49 202 - 439 1391, e-mail: *thiel@wiwi.uni-wuppertal.de*.

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). Previous empirical research that examines the link between arts activities and individual outcomes supports the subsidization of the cultural sector. Various studies report a positive relationship between cultural activities and educational attainment, cognitive skills, social capital and quality of life (e.g. DiMaggio, 1982; Hille and Schupp, 2015; Jeannotte, 2003; Kim and Kim, 2009). Despite public finding, there is however a clear education gradient in cultural engagement, and participation rates are typically higher among the better-educated (e.g. Authoring Group Educational Reporting, 2012). While this could just reflect different preferences for or experiences with arts activities, it could also involve limited access to arts goods and services due to lower incomes and lower socioeconomic status (SES), respectively. Higher social status may also lead to superior health. Those with better educational attainment and higher incomes tend to have better health outcomes (e.g. Cutler et al., 2011). Thus, the lack of financial resources to engage in social – and cultural – activities that are conducive to better health may explain social differences in health. We argue that one way to deal with these inequalities is to increase cultural participation, which is one of the major forms of the individual's social life. Many facets of individual social activities are beyond the control of policy makers, but with adequately designed public programs one could increase the consumption of cultural activities that could benefit health.

Previous research shows a positive association of cultural engagement with health status and survival. The relevance of arts activities as a therapeutic device in clinical settings is commonly acknowledged in the medical literature (McCarthy et al., 2004). More recently, several studies have stressed the role of passive or receptive arts activities for the health of the general population (e.g. Cuypers et al., 2012). To put it in a nutshell, frequent visits to cultural events may influence health directly by serving as a coping strategy or replacing health-compromising behaviors. Alternatively, cultural activities provide an enriched environment that could buffer the negative effects of age-related declines in health and cognitive skills. In addition, cultural activities can promote the establishment of social networks that provide social support for individuals facing a deterioration of their health.

However, it is debatable whether the established relationships represent causal effects of cultural participation on health. There could be unobservable variables that simultaneously influence the decision to visit a cultural event and health outcomes. For example, those who are open to experience might be more likely to attend a cultural event and have better health status. Since personality traits are usually excluded, previous analyses might overestimate the health-benefits of cultural participation. To the best of our knowledge, there is no study that seriously addresses potential confounding due to observed and unobserved characteristics and the identification of causal effects, respectively.

The aim of this study is therefore to examine whether cultural attendance has a causal effect on perceived health outcomes. This is, to the best of our knowledge, the first study regarding the impact of receptive cultural activities on health status in Germany, and we contribute to the previous literature in the following ways: First, we include a rich set of background variables including various controls for the individual's social and leisure-time activities, personality traits, health behavior and approximations for childhood exposure to the arts. These covariates were either unmeasured or excluded from previous empirical analyses, although they could influence both the decision to visit a cultural event and health status. Second, we rely on statistical matching techniques to deal with the observed and unobserved non-random selection into cultural participation. Specifically, we employ regression-adjusted propensity-score matching methods that identify causal effects given that there is no unmeasured systematic variation across individuals correlated with health outcomes. Furthermore, the matching models are combined with a difference-in-difference approach to take unobserved, time-invariant heterogeneity into account. Third, in a multivalued-treatment framework we estimate the effect of encouraging cultural activities contingent on the level of initial participation and the extent of stimulation. In other words we assess the dose-response relationship between cultural activities and health.

The empirical approach adopted in this paper can provide important insights on whether and how public programs to increase cultural participation affect health outcomes. One can think of it as a social experiment where individuals are randomly allocated to the treatment and control group. The major aim of any intervention should be to encourage individuals to regularly engage in the arts. 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. This presumably lowers the cost of consuming the arts good which should further promote cultural participation. One effective way to stimulate cultural activity would be to allocate free tickets or vouchers for cultural events among treated subjects since they most likely reflect the individual's preferences and taste (Frey, 2008). Such vouchers usually cover a relatively wide range of artistic performances, and individuals receiving the voucher can visit the event of their choice.

Hence, this study may inform policy makers about the effectiveness of using arts in general and cultural vouchers in particular to curb health inequalities. It provides a better understanding of how encouraging cultural activity affects different parts of the population which differ in terms of their current cultural behavior. The major implication of our analysis is that arts-based programs to improve population health must be targeted at those individuals who gain the most of such intervention. Our estimation results suggest that culturally inert or occasionally active individuals experience the largest health improvements whereas the health effects among more active individuals are rather negligible. Thus, distributing cultural vouchers among particularly disadvantaged parts of the population may increase the cost effectiveness of the intervention.

The remainder of this paper is organized as follows: Section (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 (3) outlines the empirical approach adopted in this paper. We generally rely on a selection-on-observables strategy using a propensity-score matching procedure. However, we will also take unobserved individual heterogeneity into account by combining the matching procedure with a difference-in-difference approach. We are thus able to control for unobserved traits that are time-invariant and possibly correlate with both the consumption of arts activities and health. Section (4) details the data and estimation sample. We use individual-level data from the German Socioeconomic Panel (SOEP) which provides rich information on, among others, the individual's social and leisure activities, personality traits, and youth socialization and activities. Section (5) presents and discusses the estimation results. Generally, visiting a cultural event frequently could improve mental health relative to lower levels of cultural activity, given a large number of covariates. The causal effect of arts attendance on mental well-being is robust to reverse causality, but disappears when we take time-fixed unobserved individual differences into account. Section (5.3) is concerned with the dose-response relationship between the health outcomes and different levels of the treatment variable. Using a multivalued-treatment approach we conduct binary comparisons of different treatment levels and control groups. Results suggest that the gain of stimulating cultural participation is highest among those respondents who less often or never visit cultural events. Section (6) concludes.

2 Previous Literature

From a theoretical perspective, there are a variety of mechanisms by 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. To maintain and accumulate their social status, individuals invest accordingly in cultural capital, for example via visits to arts events. According to 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) that social hierarchies are associated with psychosocial stress, aggressiveness, less trustfulness and lower levels of social cohesion.

Most importantly, the majority of arts activities involves social interactions with other persons that could form the basis of an individual's social capital. Individual-level social-capital indicators, 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)

Several empirical studies using individual-level survey data have found a positive association of cultural-event 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 interactions, lifestyle, personality straits, cognitive skills and childhood determinants of arts activities. Moreover, only few studies consider the possibility that health might influence cultural participation. We will improve upon the previous research by combining non-experimental evaluation methods with observational data. The problem of omitted variables bias is mitigated by including a rich set of conditioning variables that account 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 the health outcomes that are measured prior to our cultural-participation variables. Finally, we will take unobserved heterogeneity into account by using the longitudinal information of the outcome variables, and employing a difference-in-difference approach.

Complementary evidence on the causal effect of arts participation on health comes from controlled experiments in a small region of Sweden (Bygren et al., 2009; Konlaan et al., 2000). They estimated the effect of, among others, cultural-event 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 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.

The majority of the empirical studies, be it observational or experimental, relies on a comparison or control group that comprises those who never visit cultural events as well as occasionally or regularly active individuals. This might however obscure important health effects along the distribution of cultural participation in the population. As argued in Section (5.3), the health effect of increasing arts participation likely differs depending on the individual's baseline cultural activity, and less active individuals (including those who actually never participate) may benefit more than more active persons. Thus, the previous empirical evidence does little to improve our understanding of arts-based interventions to promote population health. To gain better knowledge about the workings of arts policies, such as the distribution of cultural vouchers, and their effects on health, we will also adapt a multivalued treatment framework to assess the impact of different levels of cultural participation given baseline (control) arts activity.

3 Empirical Approach

To estimate the causal parameter of interest, the empirical approach adopted in this paper relies on the *conditional-independence assumption (CIA)*. The CIA generally implies that there are no unobserved variables simultaneously affecting the treatment and the outcome. Hence, the differences in outcomes between the treated and control individuals are solely attributable to the treatment (e.g. Imbens, 2004). In other words, the treatment is assumed to be as good as randomly assigned. We will include a large set of background characteristics, which makes the CIA appear plausible in our context. However, since it is conceivable that there are still some unobserved factors causing a correlation between arts attendance and health, we will also take time-fixed unobserved individuals differences into account.

The empirical analysis begins with estimating simple linear regression models to assess the relationship between cultural attendance and health outcomes, given a set of covariates. Instead of claiming that we estimate a causal effect, we rather use it as descriptive tool to examine the observed selection bias. By successively adding groups of covariates, it can be assessed to what extent the raw correlation between arts participation and health can be explained by demographic and socioeconomic background, social and leisure-time activities, health behavior, personality, and childhood exposure to the arts.

The main part of the econometric analysis builds on matching methods to reduce the bias in observational studies. Our aim is to mimic the experimental set-up of the randomized controlled trials conducted by Konlaan et al. (2000) and Bygren et al. (2009), to create a control group that is as close as possible to the treatment group in terms of observed background characteristics. Specifically, we employ propensity-score matching methods to remove any observed pre-treatment differences between cases and controls. First, an assignment model is used to estimate the probability of receiving the treatment (e.g. frequent visits to cultural events) given the covariates. Second, the predicted values from the assignment model, the *propensity scores*, are used to match treated and control individuals using appropriate weighting functions. This implies that those control units that are more similar receive a higher weight in the estimation of the causal effect. Hence, instead of matching on combinations of covariates, the similarity of the respondents is assessed on the basis of one single number. 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. Our estimate of the causal effect is therefore assumed to be unbiased given the propensity score.

Matching methods provide results that are similar to fully saturated linear regression models (Angrist and Pischke, 2008). The empirical approach adopted in this paper has nevertheless a number of advantages: First, compared to previous studies, we include a substantially larger set of control variables in our final estimations, which should make the identifying assumptions more plausible. Second, the set of control variables is collapsed into one single number, the propensity score, which reduces the problems associated with a high-dimensional covariate vector (e.g. Huber et al., 2013). Third, matching allows assessing the comparability of treated and control individuals. Careful examination of both the covariate distributions among the treated and control units 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).

In this paper, we will estimate cross-sectional and longitudinal matching estimators with additional regression adjustment. The usefulness of combining matching with regression methods traces back to Bang and Robins (2005) and has recently been applied by Marcus (2014) and Schmitz and Westphal (2013) using SOEP data. The basic idea is to use the weights obtained from the matching procedure in a weighted regression of the outcome on the treatment and all covariates. According to Bang and Robins (2005), this ensures that the estimated causal effect is (doubly) robust against misspecifications of either the participation or the regression model. Furthermore, because the observed background characteristics are held constant, regression-adjustment reduces the bias emanating from remaining covariate imbalances after matching (Imbens and Wooldridge, 2009).

The parameter of interest is the *average treatment effect on the treated (ATT)*, that is $E\left[Y_{1i} - Y_{0i} | D = 1\right]$, with D indicating treatment status. This implies that we compare the mean health outcome of those individuals that visit cultural events (very) often with the mean health outcome had they consumed artistic performances less often. The methods employed in this paper can thus provide valuable insights into the design of public programs to encourage cultural participation and improve health. First, it can be interpreted as measuring the effect of allocating free tickets, that is vouchers, for an event per month or week among low- and non-participants. Second, we assess the health gains from such an intervention for those who actually receive the voucher and the treatment, respectively. The recipients of vouchers are probably those who are particularly disadvantaged because they lack the financial resources to afford visits to cultural events or they are unfamiliar with the potential benefits of artistic goods and services.

Since the counterfactual (Y_{0i}) is unavailable in observational studies, it is common practice to compare the outcomes of treated individuals with a weighted average of the outcomes of control matches that are as similar as possible in terms of observed characteristics. In the cross-sectional setting, the estimated ATT reads as follows (e.g. Smith and Todd, 2005):

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

where N_1 is the number individuals in the treatment group used for the calculation of the ATT, I_1 is the set of treated individuals, I_0 is the set of control individuals. $W(i,j)$ 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 will improve the inference in the cross-sectional model by including the pre-treatment (linear) trend of health outcomes among the conditioning set in a second model.

The cross-sectional matching estimator rests on the assumption that there is no confounding due to unobserved systematic differences at the individual level. It seems however likely that unobserved factors influence both the decision to attend cultural events and health outcomes. One such factor could be the individual's cognitive skills. Cognitive skills are often invoked as the main explanation for educational gradients in cultural participation and health outcomes. According to Ganzeboom (1984), individuals have different capacities to process cultural information depending on, among others, artistic talent, education and knowledge and experience associated with the arts. 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). It is possible that we only imperfectly account for the individuals' information-processing capabilities, despite including indicators for educational level and childhood exposure to the arts. Therefore, we employ the *difference-in-difference (DiD)* propensity score matching estimator which takes unobserved individual differences in outcomes into account that are time-invariant (Heckman et al., 1997, 1998). Thus, the causal effect is identified under the assumption that there are no time-varying unobserved differences leading to health-outcome differences. The DiD matching estimator is implemented by including the change scores of the dependent variables as outcomes instead of their post-treatment levels. The ATT estimator can then be written as follows (e.g. Blundell and Dias, 2009):

$$\widehat{ATT}_{DiD} = \frac{1}{N_1} \sum_{i \in I_1} \left[\Delta Y_{1i} - \sum_{j \in I_0} W(i,j) \Delta Y_{0i} \right], \tag{2}$$

where ΔY_{1i} and ΔY_{0i} denote the difference between the post- and the pre-treatment health outcome for treated and control individuals, respectively. Thus, the effect estimate measures the difference in the change of health outcomes between the treatment and control group after the treatment occurred.

In principle, various matching algorithms exist to calculate the weights $W(i,j)$, 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). We use the Epanechnikov kernel with a bandwidth equal to 0.06. The counterfactual is hence calculated as the weighted average of all control units, and more weight is given to individuals with similar propensity scores. After matching, these weights are used in a weighted least squares regression to obtain the ATT. Since the choice of the kernel bandwidth involves a trade-off between bias and efficiency, we will also estimate the matching models using a bandwidth of 0.03. Additionally, we will employ the kernel matching estimator using the Gaussian kernel, and other matching algorithms based on radius and nearest-neighbor methods, as additional robustness checks.

As argued above, it is crucial for the estimation of average treatment effects that one compares treated and control individuals that are similar. Hence, 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. According to this rule, we drop 19 to 21 individuals whose propensity score is too high, depending on the matching model.

Furthermore, the treatment and control group must 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 (2014), 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. 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.

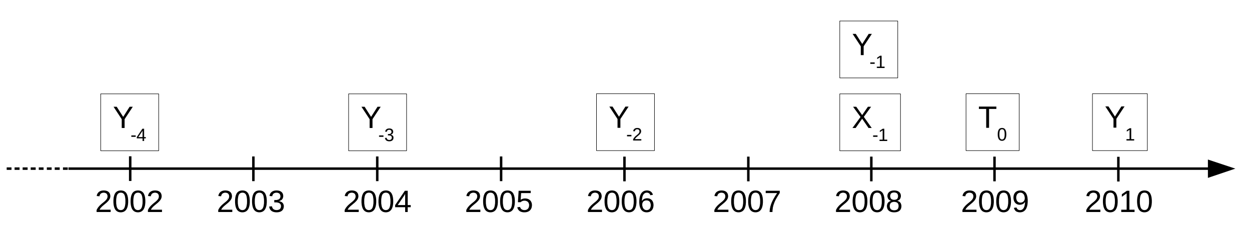
A final issue concerns the estimation of the variance of the treatment effects. Generally, the estimated standard errors neglect the estimation of the propensity score. The uncertainty associated with the propensity-score estimation is thus disregarded, which could increase the variance of the treatment effects (Heckman et al., 1998). We therefore follow Marcus (2014) and Schmitz and Westphal (2013) 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.

4 Data

4.1 Description of the estimation sample

This study uses longitudinal data from the Socio-Economic Panel (SOEP) study (SOEP, 2013; Wagner et al., 2007). It is a large and representative survey of German households which started in 1984. It is well-suited for our purposes since it includes information on health status, demographic and socioeconomic background, leisure-time and social activities, personality and youth socialization. Figure (1) illustrates the basic structure of the estimation sample. The treatment variable T_0 is measured in 2009. While information on 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-time and social activities around this year. The health outcomes Y_1 are measured in 2010. We will also consider the health outcomes measured in 2012 to assess the longer-term impact of cultural activities. Furthermore, we will use information on health outcomes prior to 2009 to mitigate the problem of reverse causality, that is the possibility that the outcome influences the treatment. Most of the control variables X are gathered in 2008 prior to the treatment, to alleviate the problem that some (time-varying) background characteristics are influenced by the treatment or the anticipation of it. We will also consider health outcomes assessed prior to the treatment variable (Y_{-4} to Y_{-1}) to account for the bi-directional relationship between cultural participation and health.

Figure 1: Structure of the estimation sample



The final analysis sample includes 4,267 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.

4.2 Definition of the treatment variable

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. The treatment variable is calculated 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 "at least once a week", "at least once a month", "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 (Frey, 2008). To make our results comparable to previous experimental evidence (Bygren et al., 2009; Konlaan et al., 2000), the treatment 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 that visit cultural events at least once a month ($n = 953$). The control group comprises those respondents that less often or never visit cultural events ($n = 3,314$). This is akin to randomly assigning treatment status among those who rarely or never engage in cultural activities, and giving vouchers for a free cultural event per week or month to subjects in the treatment group.

As already argued in Section (2), defining the control group in this way leaves us uninformed about the effect of promoting cultural participation for individuals with different baseline states of cultural activity. In Section (5.3), we therefore conduct pairwise comparisons between different treatment and control levels of cultural participation. This allows us to assess how health reacts to different levels of cultural stimulation given baseline (control) arts-participation status.

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 questionnaire and a widely accepted and validated instrument 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*). 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 (48.3) than the general population, but the difference seems rather negligible (see Table (A.1) in the appendix).[1]

4.4 Definition of the conditioning set

For the unconfoundedness assumption to be plausible and to identify a causal effect, a matching analysis must include all variables that influence both the treatment and the outcome (e.g. Caliendo and Kopeinig, 2008). The decision to attend cultural events is non-random and can be couched in terms of a constrained utility-maximization problem (e.g. Gray, 2011). 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. 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/social capital, leisure-time activities, health behavior, personality traits, and childhood exposure to the arts (see Table (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

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

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 artistic or health-related activities. The impact of the urbanization level possibly reflects 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)

As discussed above, the main mechanism resulting in a positive arts-health relationship are the social interactions usually involved in cultural activities. We therefore 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-time 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. Furthermore, creative activities have been shown to be correlated with perceived health (e.g. Cuypers et al., 2012). On the other hand, other leisure-time activities could reflect time constraints that reduce the opportunities to engage in both cultural and health enhancing activities.

Health-related lifestyles could also influence arts participation and perceived health. Therefore, we control for health behaviors such as body-mass index (BMI), smoking status, alcohol intake and dietary behavior. Furthermore, we seriously acknowledge the possibility of reverse causality by conditioning on the pre-treatment (linear) trend of the health outcome variables.

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 a 4-year period (see also Costa and McCrae, 1988).

As argued above, current arts consumption or participation is a function of past arts exposure and consumption. Previous research basically suggest two pathways through which past behavior influences current arts-participation levels. From a rational-choice perspective, consumption of arts goods and services can have addictive patterns. Visiting cultural events or attending music classes can be seen as investments in arts-specific consumption capital that lower the cost of subsequent cultural consumption, potentially raising future demand for cultural goods (Stigler and Becker, 1977) A different approach assumes that tastes are unknown and revealed to the consumer when experiencing artistic performances (Lévy-Garboua and Montmarquette, 1996). 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. Hence, past consumption of arts goods likely influences current arts activities through the formation of tastes and preferences for arts appreciation. This corresponds to the conjecture that arts consumption has the properties of an experience good. Individuals are often unaware of the (health) benefits of arts activities, and they learn to appreciate the cultural good only after they have consumed it (e.g. Lévy-Garboua and Montmarquette, 2011).

To approximate early influences on adult health and arts participation, 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, and whether the individual did sports and attended musical lessons during youth. The empirical models in previous studies, to the best of our knowledge, lack these factors although at least theoretically they may account for a large part of the variation in adult cultural participation (e.g. Morrison and West, 1986).

5 Empirical Results

5.1 Regression estimates of the effect of cultural participation

We first present results from simple linear regression models of the outcomes, where the set of covariates is successively included. If we multiply the coefficients with ten, the estimates can also be interpreted in percent of a standard deviation. The first column of Table (1) reports the unconditional (raw) association between the treatment 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 binary treatment indicator increases the physical and mental health scores by 1.6 and 2.8 points (or 16 and 28 percent of a standard deviation), respectively. This relationship is significant at the 99 percent level.

Table 1: OLS results, different specifications

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Outcomes								
Physical health	1.558***	2.506***	1.262***	1.129***	0.837**	0.640*	0.546	0.547
	(0.355)	(0.321)	(0.329)	(0.337)	(0.354)	(0.349)	(0.342)	(0.342)
Mental health	2.839***	2.207***	1.708***	1.383***	1.178***	1.094***	0.780**	0.800**
	(0.323)	(0.329)	(0.355)	(0.361)	(0.375)	(0.375)	(0.350)	(0.350)
Control variables								
Demographic background		✓	✓	✓	✓	✓	✓	✓
Soecioeconomic status			✓	✓	✓	✓	✓	✓
Social activities				✓	✓	✓	✓	✓
Leisure-time activities					✓	✓	✓	✓
Health behavior						✓	✓	✓
Personality traits (Big 5)							✓	✓
Childhood conditions								✓

$N = 4,267$. Robust standard errors in parentheses. * $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$.

In general, including the covariate groups gradually decreases the marginal effect of cultural attendance on 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 the treatment based on observable characteristics. The second column of Table (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. 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-event visits.[2]

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 overestimate the effect of attendance on health.

The coefficients reported in the fourth column suggest that part of the arts-health relationship can be explained by social activities or the individual's social capital. 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, particular in the case of physical health.

Including leisure-time activities and health behaviors in the fifth and sixth column further decreases the effect of 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, previous studies likely overestimate the effect of attending cultural events on health because they usually do no condition on leisure-time or health-related activities.

The seventh column of Table (1) shows, that there is some selection based on personality traits. On the one hand, more open individuals are possibly more likely to attend cultural events and exhibit less health problems. On the other hand, respondents with high scores in

[2]These explanations are supported by the positive association of age with cultural attendance and the negative (positive) correlation between attendance and physical health (mental health) (results not shown here).

neuroticism or low scores in emotional stability are probably less likely to engage in cultural participation and have worse health outcomes. Both factors tend to produce overestimates of the effect of cultural participation on health when we exclude them.

The last column reports the marginal effects of cultural participation on health conditioning on all covariates including the approximations for childhood exposure to the arts. However, the coefficients for the treatment 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 arts-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.

5.2 Matching estimates of the effect of cultural participation

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.[3] Age positively influences the probability of attending cultural events at least monthly, 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.

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 house-

[3]See Appendix (B) for full results.

hold 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 to pursue an active lifestyle, socialize more, and exhibit higher levels of political and civic engagement. What is more, cultural-event 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 (3) is employed. The aim is to find control individuals that are similar to treated subjects in terms of their covariate distribution. Hence, it is a critical task in any matching analysis to assess the covariate balance between the treatment and control group. This is usually done by evaluating the standardized difference in each independent variable. We will do this for each covariate and present the median value of all standardized difference for each matching model.

To illustrate this step, Table (A.3) in the appendix reports mean values and the standardized bias for each covariate in the treatment and control group, and before and after matching, for the cross-sectional matching estimator. 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 11 percent after matching. This is however 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) emanate from a regression of the propensity score on the covariates using the unmatched (raw) and matched sample, respectively. They suggest that the explanatory power of the regressors is fairly low in the matched sample (0.01) compared to the unmatched sample (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 2.2 while the median bias amounts to 1.6, reflecting a considerable reduction in terms of covariate imbalance compared to the unmatched sample (13.4 and 6.8).

Table 2: Overall statistics on covariate balance

Sample	Pseudo R2	LR chi2	p>chi2	Mean bias	Median bias
Unmatched	0.30	1338.38	0	13.4	6.8
Matched	0.01	34.69	1	2.2	1.6

Causal-effect estimates. Table (3) provides causal estimates of the effect of cultural-event attendance on health using regression-adjusted propensity-score matching models. The coefficients in each cell represent average treatment effects on the treated (ATT). The figures in parentheses are robust standard errors from weighted linear regressions of the outcome on the treatment variable and covariates, which omit the estimation of the propensity score. We also computed bootstrapped standard errors as a robustness check, but the results basically remain the same.

The first column of Table (3) displays the $ATTs$ from what can be called the simple cross-sectional matching estimator, where we exclude longitudinal information on outcome variables. The estimates are fairly similar to the linear regression results. The ATT for physical health is equal to 0.43, which is slightly smaller than the OLS estimate (0.55). The effect is, however, not significant, which can mainly be attributed to the reduced effect estimate. The ATT for mental health amounts to 0.91, which is slighty higher than the OLS estimate. The effect is still significant at the 95 percent level. In other words, had the treated respondent not received the cultural voucher, his or her mental health score would have been 0.91 points lower (or about 9 percent of a standard deviation).

The slight discrepancy between the matching and linear regression results can be explained by the different weighting schemes they apply. According to Angrist and Pischke

Table 3: Matching estimates using 2010 outcomes

	(1) Cross-sectional	(2) + pre-treatment health trends	(3) DiD
Physical health			
ATT	0.433	0.453	-0.165
	(0.369)	(0.366)	(0.306)
N_{Treated}	932	932	932
N_{Controls}	3,283	3,283	3,283
$N_{\text{Off support}}$	21	21	21
Median bias (%)	1.6	1.7	1.7
Mental health			
ATT	0.917**	0.982**	0.396
	(0.409)	(0.419)	(0.399)
N_{Treated}	932	934	934
N_{Controls}	3,283	3,283	3,283
$N_{\text{Off support}}$	21	19	19
Median bias (%)	1.6	1.5	1.5

Robust standard errors in parentheses. * $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$. All models condition on demographic background, socioeconomic status, social and leisure-time activities, health behavior, personality traits, and childhood exposure to the arts.

(2008), matching puts more weight on control individuals being more likely to be treated. On the one hand, the matching estimates are smaller because these individuals also have less physical health problems per se. On the other hand, matching estimates are larger than OLS estimates because control individuals with a high probability of receiving the treatment are probably more likely to exhibit mental health problems.

The second column of Table (3) reports ATT estimates including pre-treatment health outcomes among the conditioning variables. Specifically, we include the pre-treatments (linear) trend in health outcomes (between 2002 and 2008) to take the problem of reverse causality into account. The results suggest that there is some health-related selection into cultural participation with respect to mental health only. The ATT estimate increases somewhat from 0.91 to 0.98, and is still significant at the 95 percent level. Thus, if reverse causality is an issue, it likely results in an underestimate of the causal effect on mental health.

The third column of Table (3) reports causal-effect estimates from the regression-adjusted propensity-score matching estimator, combined with a difference-in-difference approach. Hence, we include the change scores between post-treatment and pre-treatment health as outcome variables, and include all control variables. Inference in difference-in-difference models hinges on the common trend assumption, that is the same trend of outcome variables in the absence of treatment. A direct test of this assumption is virtually impossible since the counterfactual is unavailable. However, inspection of the pre-treatment data suggests that the pre-treatment trends in health outcomes in both the treatment and the control group are basically the same (see Figure (A. 1) in the appendix).

The difference-in-difference approach combined with a matching procedure yields a more credible version of the CIA. Thus, the coefficients can be interpreted as a causal effect conditional on observable characteristics and unobserved variables that are time-invariant. As discussed above, cognitive skills are correlated with both cultural participation and health, but are usually unmeasured in large-scale observational surveys. Taking time-fixed unobserved confounders into account considerably reduces the mental-health effects of cultural attendance (from 0.98 to 0.4). Hence, there are unobserved characteristics, such as cognitive skills or talent, that produce positive mental-health benefits of cultural participation. Making the treated and control individuals equal in terms of observable and time-invariant unmeasured covariates thus eliminates any difference in health outcomes that could be attributable to arts participation. This suggests that stimulating engagement in cultural participation by offering free tickets for an event per month or week is rather ineffective in improving individual health status.

This corresponds to some extent with the results reported by Bygren et al. (2009), who did not found a significant impact of cultural attendance on similar measures of physical and mental health. As argued in section (5.3), however, comparing individuals that frequently visit cultural events (at least once a month) with others who rarely or never consume artistic goods and services may obscure the health benefits of cultural participation. Specifically, whether individuals gain from greater engagement in culture and arts possibly depends on the extent of cultural stimulation and their baseline participation status, and the definition of the treatment level and the control outcome, respectively. We will elaborate further on this issue in Section (5.3), by assigning respondents different treatment levels.

Longer-term impact and robustness checks.[4] We also estimated the regression-adjusted propensity-score matching models using health outcomes measured in 2012. The effect of cultural-event attendance on both physical and mental health is statistically insignificant in the cross-sectional and difference-in-difference models. Thus, cultural participation seems to have no lasting impact on individual health status. This is even true if we ignore time-invariant unobserved heterogeneity.

[4]For detailed results, see Appendix (B) and (C).

A number of robustness checks suggest that the main findings of the matching analyses are rather insensitive to various model specifications. We obtain similar ATT estimates using different matching algorithms (lower bandwidth, Gaussian kernel, nearest-neighbor matching), different sets of conditioning variables (only significant covariates, excluding potentially endogenous variables), and bootstrapped standard errors.

5.3 Multivalued treatment effects

Up to now we have compared individuals that visited cultural events at least once a month with those who attended cultural events less often. The difference-in-difference results suggest that the former are just as healthy as the latter. Defining the treatment and control group in this way, however, could mask health differences across the different levels of cultural participation. For example, even rare visits to cultural events may provide health benefits relative to non-participation, while increasing cultural participation from at least once a month to at least once a week might be unrelated to health outcomes.

The effectiveness of cultural-event attendance as a means to improve health outcomes therefore hinges on the extent of stimulation, that is whether individuals are encouraged to attend cultural performances at least once a week, at least once a month or less often. What is more, individuals might be differently affected by these treatments, depending on their initial level of cultural activity (e.g. Cattaneo, 2010). A better understanding of heterogeneous effects across treatment levels is also important from a policy perspective. It ensures that economic resources are geared towards those individuals that gain the most from cultural participation. The results based on the collapsed binary indicator suggest that frequent cultural participation, relative to fewer activities, is unimportant for health outcomes. Hence, from a cost-effectiveness perspective it appears economically unsound to encourage cultural engagement by allocating vouchers for an event per month or week among low or non-consumers. Nonetheless, this still leaves the door open for the possibility that incremental increases in cultural participation are beneficial. Suppose, for instance, that one could credibly ascertain a positive effect of occasional cultural attendance on health relative to non-participation. This would suggest that policy makers can improve the health among the non-participants with relatively little effort.

In this section, we use a multivalued treatment framework and compute the dose-response relationship between health outcomes and different levels of cultural activity. Specifically, we use all four categories (at least once a week, at least once a month, less often, never) of the cultural-attendance variable. Instead of adopting an ordered or multinomial choice model for the participation equation, we conduct pairwise comparisons and estimate a series of binomial models. According to Lechner (2002), this approach provides similar results and is less prone to misspecification than multinomial or ordered choice models. The matching algorithm is then applied to each of the following pairwise comparison: *at least once a week* vs. *at least once a month*, *at least once a week* vs. *less often*, *at least once a week* vs. *never*, *at least once a month* vs. *less often*, *at least once a month* vs. *never*, and *less often* vs. *never*.

Thus, the participation probabilities and causal-effect estimates are based on the subpopulation confined to individuals that belong to either of the two groups. For each pairwise comparison, we estimate the three propensity-score kernel-matching estimators (cross-sectional, including pre-treatment health trends, difference-in-difference) presented above. As shown by Imbens (2000) and Lechner (2002), the interpretation of $ATTs$ in a multivalued treatment framework as causal effects rests on the same identifiying assumptions (conditional independence and overlap) as in the binary case.

Figure (2) displays the (unconditional) estimated mean of the physical and mental health scores for each level of the cultural-attendance variable. Generally, health improves with the frequency of cultural-event attendance. On the one hand, individuals who visited cultural events less often, at least once a month, or at least once week exhibit greater physical-health scores than those who never attended cultural events. As indicated by the overlapping confidence intervals, however, it appears that physical health is basically the same among those who at least occasionally visit cultural events. On the other hand, the dose-response relationship seems to be more pronounced for mental health. The mental-health score seems to be continuously increasing in the level of cultural participation, and the relationship flattens only for the two highest categories.

Table (4) shows the $ATTs$ from the difference-in-difference propensity score models for each pairwise comparison.[5] The rows include the treatment levels and the columns represent the control groups, respectively. For example, the physical health score decreases by 0.4 points when cultural participation increases from very low (never) to slightly higher (less often) levels (not significant).

It appears that cultural-event attendance is particularly beneficial for non-participants. Their mental-health score rises by 1.5 points when they increase cultural participation to "less often", while the corresponding improvement induced by increasing cultural participation to "at least once a month" and "at least once a week" equals 3 and 4.7 points, respectively. These estimates are significant at the 99 percent level. The standardized covariate difference between treated and control units is comparatively large, with a median bias amounting

[5]Full results including the cross-sectional matching estimator and the matching model including pre-treatment health outcomes are shown in Table (E.1) in the appendix.

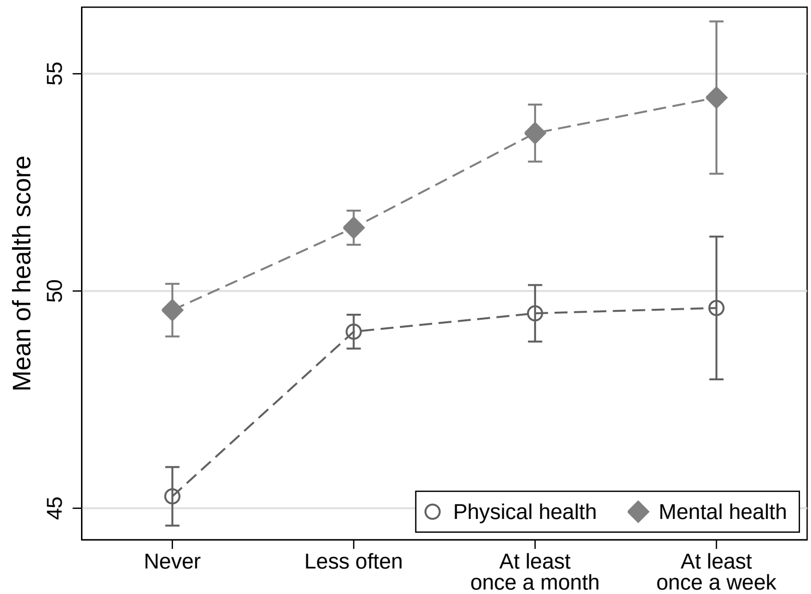


Figure 2: Health outcomes by level of cultural participation

to no less than 11% and 6%, respectively. The comparison of health outcomes between those who "less often" visit cultural events and those who "never" participate apparently provides a more credible causal effect. The observed differences between treated and control individuals are relatively low (2.4%), indicating that the treatment and control group are alike. Hence, even low levels of cultural participation may yield mental-health benefits for those who never participate.

Nonetheless, the mental-health effects of cultural participation seem to vanish for those who regularly attend cultural events. Increasing cultural participation from "less often" to "at least once a month" or from "at least once a month" to "at least once a week" is unrelated to mental health. Those who visit cultural events less often only experience a mental-health increment if they increase their cultural participation to a relatively large extent. Their mental-health score rises by 1.8 points when they visit cultural events at least once a week (significant at the 95 percent level).

6 Conclusion

In this paper, we examined the causal effect of cultural-event attendance on perceived health outcomes by seriously taking non-random selection due to observable and unobservable characteristics into account. Results from simple linear regression models suggest that the correlation between arts attendance and mental and physical health is potentially confounded by the

Table 4: Differential effects of treatment levels

	Control group		
Treatment group	Never	Less often	More than once a month
Physical health			
Never			
Less often	-0.072		
	(0.346)		
More than once a month	-0.418	-0.095	
	(0.434)	(0.318)	
More than once a week	-0.767	0.015	0.423
	(1.418)	(0.533)	(0.600)
Mental health			
Never			
Less often	1.487***		
	(0.461)		
More than once a month	3.072***	-0.037	
	(0.598)	(0.380)	
More than once a week	4.747***	1.750**	0.638
	(1.311)	(0.696)	(0.604)

Estimated ATTs come from regression-adjusted, difference-in-difference propensity-score matching models for each pairwise comparison. Robust standard errors in parentheses. * $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$. All models condition on demographic background, socioeconomic status, social and leisure-time activities, health behavior, personality traits, childhood exposure to the arts, and pre-treatment health outcomes.

respondent's demographic background, socioeconomic status, social and leisure-time activities, personality traits and health behavior. Estimates from regression-adjusted propensity-score matching models reveal that attending arts activities at least once a month can improve mental well-being compared to less frequent visits. This causal effect is robust to reverse causality and health-related selection, respectively. However, results from difference-in-difference matching models suggest that the causal-effect estimate can be explained by time-invariant unobserved heterogeneity. In our context, a possible candidate responsible for unobserved, time-fixed differences between individuals are cognitive or information-processing skills. Although we control and approximate for educational attainment, personality, and childhood exposure to the arts, it is possible that we only imperfectly or incompletely account for the respondent's cognitive skills in the cross-sectional matching models. A number of binary comparisons between different treatment levels and control groups reveal that the health benefits of cultural participation can be substantial, particularly among those who rarely or never attend. Low levels of cultural participation can therefore improve the mental-health outcomes even of those individuals who never visit artistic activities. This result seems plausible since receptive cultural activities usually involve only limited physical activity and are more likely to affect psychological well-being through, for instance, an enriched environment and the health effects of social networks. Furthermore, inactive individuals or those who have only little experience with cultural activities appear to receive the greatest health improvements from increased cultural participation.

This has important implications for public policy measures that use arts to improve health outcomes. Our results suggest that stimulating cultural activity by allocating free tickets or vouchers for only few events per year among non and low consumers of cultural services might improve the health status of these individuals to a large extent. Specifically, the vouchers can be used to initiate cultural engagement. Individuals that rarely or never attend arts activities probably have no or only little previous experience with arts goods and services, and are therefore unaware of their potential benefits. The value of these goods, similar to that of experience goods, is appreciated only after they have consumed it. By using the voucher, the individuals may learn to appreciate the artistic good and the cost of arts consumption declines, which presumably increases future cultural activity. Hence, a program that distributes vouchers for cultural events among those who rarely or never attend may represent an economically feasible way to improve the mental well-being of those who are incapable of engaging in cultural activity. It would furthermore suffice that such an intervention is restricted to a limited period of time. The users of the voucher, if they learn the benefits of and appreciate cultural activities, may have an economic incentive to visit cultural events in the future even in the absence of the voucher. Thus, cultural vouchers can be a cost-effective means to reduce social inequalities in mental health by facilitating access to cultural events for particularly disadvantaged persons.

However, the empirical strategy employed in this paper has several limitations that could undermine inference and policy implications. First, we use a rather general indicator of (high-brow) cultural attendance, which aggregates different types of events into one single variable. It could be possible that different types of events have different impacts, and that these effects might cancel out the health benefits. Second, we were unable to ascertain the health benefits of arts-related programs that aim to communicate and promote health messages. Hence, our results are confined to the physical and psychological consequences of general cultural participation with no explicit reference to health. Third, there could exist other unobserved factors that vary with time, and could influence both cultural participation and health outcomes. Therefore, it is possible that the causal-effect estimates are biased and that we still over- or underestimate the health benefits of arts participation. One way to deal with this type of endogeneity would be to apply an instrumental variable approach. This involves finding exogenous variation in cultural attendance which is unrelated to the outcomes via other unobserved factors. Potential candidates would be local supply-side indicators that reflect the availability and accessibility of cultural institutions.

Bearing the shortcomings of causal inference with observational data in mind, the "gold standard" to ascertain the causal effect of cultural participation on health are still experimental designs. Indeed, the research question at hand lends itself to a social experiment that could involve, for instance, the distribution of cultural vouchers among the inactive or occasionally active population. Random assignment would ensure that the composition of the treatment and control group is the same and that any health difference between these two groups can be attributed to the voucher. Assignment of treatment status would be under the control of the researcher implying a more credible version of the unconfoundedness assumption as with observational data. We would obtain causal-effect estimates of arts participation that are presumably unbiased and less vulnerable to systematic selection into the treatment arising from unobserved heterogeneity and simultaneity. Future research should therefore include a social experiment that could complement and substantiate our non-experimental findings for Germany.

Acknowledgements

I gratefully acknowledge valuable comments and suggestions from Hendrik Jürges. I also thank participants for fruitful discussions in the research seminars at the Schumpeter School of Business and Economics in Wuppertal and at the annual congress of the German Health Economics Association.

References

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.

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.

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* **127**(1): 171–182.

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.

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

Blundell, R. and Dias, M. C. (2009). Alternative approaches to evaluation in empirical microeconomics, *Journal of Human Resources* **44**(3): 565–640.

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

Brownstone, D. and Valletta, R. (2001). The bootstrap and multiple imputations: harnessing increased computing power for improved statistical tests, *Journal of Economic Perspectives* **15**(4): 129–141.

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.

Cattaneo, M. D. (2010). Efficient semiparametric estimation of multi-valued treatment effects under ignorability, *Journal of Econometrics* **155**(2): 138–154.

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.

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.

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.

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.

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.

DiNardo, J. and Tobias, J. L. (2001). Nonparametric density and regression estimation, *Journal of Economic Perspectives* **15**(4): 11–28.

Folland, S. (2008). An economic model of social capital and health, *Health Economics, Policy and Law* **3**(4): 333–348.

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.

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.

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.

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

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

Heckman, J. J., Ichimura, H. and Todd, P. E. (1997). Matching as an econometric evaluation estimator: Evidence from evaluating a job training programme, *Review of Economic Studies* **64**(4): 605–654.

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**(0): 56–82.

Huber, M., Lechner, M. and Wunsch, C. (2013). The performance of estimators based on the propensity score, *Journal of Econometrics* **175**(1): 1–21.

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.

Imbens, G. (2014). Matching methods in practice: Three examples, *NBER Working Paper No. 19959*, National Bureau of Economic Research.

Imbens, G. W. (2000). The role of the propensity score in estimating dose-response functions, *Biometrika* **87**(3): 706–710.

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. 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.

Jeannotte, M. S. (2003). Singing alone? The contribution of cultural capital to social cohesion and sustainable communities, *The 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.

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.

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.

Lechner, M. (2002). Program heterogeneity and propensity score matching: An application to the evaluation of active labor market policies, *Review of Economics and Statistics* **84**(2): 205–220.

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

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

Lévy-Garboua, L. and Montmarquette, C. (2011). Demand, *in* R. Towse (ed.), *A Handbook of Cultural Economics*, 2nd edn, Edward Elgar, Cheltenham, pp. 177–187.

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

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

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.

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.

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, *The American Statistician* **39**(1): 33–38.

Schmitz, H. and Westphal, M. (2013). Short-and medium-term effects of informal care provision on health, *SOEPpapers No. 569*, DIW, Berlin.

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.

Socio-Economic Panel (SOEP) (2013). Data for years 1984-2012, Version 29, SOEP. doi: doi:10.5684/soep.v29.

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, *The Journals of Gerontology Series B: Psychological Sciences and Social Sciences* **62**(Special Issue 1): 62–69.

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.

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.

Appendix A Tables and Figures

Table A.1: Summary statistics ($N = 4,267$)

Variable	Definition	Mean	SD	Min.	Max.
Treatment variable					
Cultural-event attendance	1= visits cultural events at least once a month or week, 0=otherwis	0.22	0.42	0	1
Outcome variables					
Mental Health	Physical component summary scale (pcs) from SF-12 questionnaire	51.53	9.77	7.65	74.18
Physical Health	Mental component summary scale (mcs) from SF-12 questionnaire	48.29	10.00	15.02	73.08
Demographic background					
Sex	1=female, 0= male	0.52	0.50	0	1
Age	Age of the individual (included as dummy variables)	55.53	14.63	24	99
Number of persons	Log. of number of persons in the household	0.81	0.47	0	1.95
Number of children	Log. of children under the age of 16 in the household	0.23	0.42	0	1.79
Married	1=married or living together, 0= otherwise	0.70	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
Region undergoing urbanization	1=living in region undergoing urbanization, 0=otherwise	0.30	0.46	0	1
Rural region	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

Continued on next page...

... table A.1 continued

Variable	Definition	Mean	SD	Min.	Max.
Intermediate track	1=intermediate school degree, 0=otherwise	0.28	0.45	0	1
Academic track	1=leaving certificate from vocational school or college entry exam, 0=otherwise	0.35	0.48	0	1
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. of net equivalent household income	10.02	0.60	0	12.84
Employed	1=employed full- or part-time, 0=otherwise	0.58	0.49	0	1
Not employed	1=not employed, 0=otherwise	0.38	0.49	0	1
Unemployed	1=registered unemployed, 0=otherwise	0.04	0.20	0	1
Social capital/activities					
Volunteer engagement	1=volunteer at least once a month or week, 0=otherwise	0.22	0.42	0	1
Political engagement	1=participate in political activities at least once a month or week, 0=otherwise	0.04	0.19	0	1
Religious participation	1=attend church or religious events at least once a month or week, 0=otherwise	0.19	0.40	0	1
Neighbors and friends	1=visit neighbors and friends at least once a month or week, 0=otherwise	0.76	0.43	0	1
Relatives	1=visit relatives at least once a month or week, 0=otherwise	0.76	0.43	0	1
Leisure-time activities					
Cinema, pop concerts, disco	1=visit cinemas, pop concerts or discos at least once a month or week, 0=otherwise	0.15	0.36	0	1
Sports	1=exercise at least once a week, 0=otherwise	0.44	0.50	0	1

Continued on next page...

... table A.1 continued

Variable	Definition	Mean	SD	Min.	Max.
Eating, drinking	1=go out for a meal or drink at least once a month or week, 0=otherwise	0.55	0.50	0	1
Excursions, short trips	1=go on excursions or trips at least once a month or week, 0=otherwise	0.29	0.45	0	1
TV, video	1=watch TV or video daily, 0=otherwise	0.83	0.37	0	1
Computer use	1=use computer privately at least once a week, 0=otherwise	0.58	0.49	0	1
Artistic activities	1=pursue artistic activities at least once a month or week, 0=otherwise	0.24	0.43	0	1
Garden work	1=do garden work, hand crafts or repairing at least once a month or week, 0=otherwise	0.67	0.47	0	1
Car repair/maintenance	1=do car repair or maintenance at least once a month or week, 0=otherwise	0.25	0.43	0	1
Sport events	1=attend sport events at least once a month or week, 0=otherwise	0.11	0.32	0	1
Health behavior					
BMI	Body mass index calculated as (weight in kgs/height in ms)	26.34	4.62	16.10	67.20
Smoking status	1=smoker,0=otherwise	0.23	0.42	0	1
Alcohol intake	1=drink alcohol regularly, 0=otherwise	0.22	0.42	0	1
Nutrition	1=keep healthy diet, 0=otherwise	0.43	0.50	0	1
Personality traits (2005 and 2009 average)					
Conscientiousness	Score on the conscientiousness scale	5.96	0.79	2	7
Openness	Score on the openness scale	4.55	1.07	1	7
Extraversion	Score on the extraversion scale	4.82	1.02	1.33	7
Agreeableness	Score on the agreeableness scale	5.43	0.86	1.83	7
Neuroticism	Score on the neuroticism scale	3.77	1.11	1	7
Childhood conditions					
Sports activities during youth	1=did sports during youth, 0=otherwise	0.56	0.50	0	1

Continued on next page...

... table A.1 continued

Variable	**Definition**	**Mean**	**SD**	**Min.**	**Max.**
Musical activity in youth	1=played and instrument during youth, 0=otherwise	0.32	0.47	0	1
Basic track (parents)	1=secondary general school leaving certificate or no degree, 0=otherwise	0.73	0.44	0	1
Intermediate track (parents)	1=intermediate school degree, 0=otherwise	0.18	0.38	0	1
Academic track (parents)	1=leaving certificate from vocational school or college entry exam, 0=otherwise	0.08	0.27	0	1
Training (parents)	1=vocational degree, o=otherwise	0.55	0.50	0	1
University degree (parents)	1=university degree, 0=otherwise	0.06	0.23	0	1
Large city	1=lived in large city during childhood, 0=otherwise	0.25	0.43	0	1
Medium city	1=lived in medium city during childhood, 0=otherwise	0.18	0.39	0	1
Small city	1=lived in small city during childhood, 0=otherwise	0.22	0.41	0	1
Countryside	1=lived in the countryside during childhood, 0=otherwise	0.35	0.48	0	1
Siblings	Number of the respondent's siblings	1.90	1.69	0	13

Table A.2: SOEP health scales

Subscale (# of items)	Response scale/Questionnaire wording
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?

Continued on next page...

... table A.2 continued

Subscale (# of items)	Response scale/Questionnaire wording
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

Table A.3: Covariate balance statistics

	Unmatched			Matched		
Variable	Treated	Control	Bias	Treated	Control	Bias
Sex	0.52	0.52	0.9	0.52	0.54	-3
Number of persons	0.75	0.83	-17.6	0.76	0.78	-4.7
Number of children	0.15	0.25	-25.1	0.15	0.18	-6.1
Separated	0.02	0.02	-0.4	0.02	0.02	-1.3
Single	0.14	0.14	-0.5	0.14	0.15	-2.5
Divorced	0.08	0.08	0.7	0.08	0.08	-0.6
Widowed	0.06	0.07	-4.4	0.06	0.06	-2
Urban region	0.57	0.46	23	0.57	0.59	-3.6
Region undergoing urbanization	0.24	0.31	-15.5	0.25	0.24	2.3
Intermediate track	0.24	0.29	-11.9	0.24	0.27	-6.9
Academic track	0.57	0.29	59.2	0.56	0.54	4.2
Training	0.57	0.77	-43.6	0.58	0.59	-3.3
University degree	0.55	0.25	65.9	0.54	0.53	3.3
Household income	10.32	9.94	65.2	10.31	10.30	1.6
Not employed	0.44	0.35	16.5	0.43	0.40	6.2
Unemployed	0.03	0.05	-10.9	0.03	0.03	0
Volunteer engagement	0.34	0.19	32.8	0.33	0.34	-3.3
Political commitment	0.10	0.02	32.5	0.09	0.09	-1.8
Religious participation	0.27	0.17	22.7	0.26	0.25	1.1
Neighbors and friends	0.86	0.73	33.8	0.86	0.85	1.9
Relatives	0.76	0.76	-0.1	0.76	0.76	-0.6
Cinema, pop concerst, disco	0.31	0.11	51.5	0.30	0.30	0.1
Sports	0.63	0.39	49.8	0.63	0.63	-0.7
Eating, drinking	0.78	0.49	64.1	0.78	0.78	0.6
Excursions, short trips	0.49	0.23	57.3	0.48	0.49	-0.1
TV, video	0.78	0.85	-18	0.78	0.77	2.1
Computer use	0.68	0.55	27.8	0.68	0.70	-3.9
Artistic activities	0.45	0.18	61.3	0.44	0.44	-1
Garden work	0.69	0.67	5.2	0.69	0.68	2.8
Car repair/maintenance	0.20	0.27	-16.4	0.20	0.21	-3
Sport events	0.13	0.11	8.5	0.13	0.14	-1.1
BMI	25.52	26.59	-24.3	25.52	25.40	2.7
Smoking status	0.15	0.26	-27	0.15	0.16	-2.7

Continued on next page...

... table A.3 continued

Variable	Unmatched			Matched		
	Treated	Control	Bias	Treated	Control	Bias
Alcohol intake	0.28	0.20	18.1	0.28	0.28	-0.3
Nutrition	0.30	0.47	-34.4	0.31	0.30	1
Conscientiousness	5.91	5.97	-7.1	5.91	5.91	-1
Openness	5.00	4.42	56.9	4.98	5.03	-4.8
Extraversion	4.94	4.79	15.5	4.93	4.94	-1.1
Agreeableness	5.48	5.41	7.8	5.47	5.48	-1
Neuroticism	3.59	3.82	-21	3.60	3.63	-2.5
Musical activity in youth	0.46	0.28	35.9	0.45	0.47	-3.8
Sports activities in youth	0.61	0.54	14.4	0.61	0.59	3
Intermediate track (parents)	0.24	0.16	19.9	0.24	0.23	1.4
Academic track (parents)	0.13	0.06	24.7	0.13	0.14	-5.4
Training (parents)	0.57	0.55	4.5	0.57	0.58	-1
University degree (parents)	0.08	0.05	13.8	0.08	0.11	-11
Medium city	0.20	0.18	5	0.19	0.19	1.1
Small city	0.20	0.22	-3.1	0.21	0.21	-0.5
Countryside	0.29	0.37	-18.6	0.29	0.28	1
Siblings	1.72	1.95	-13.9	1.72	1.74	-1.7

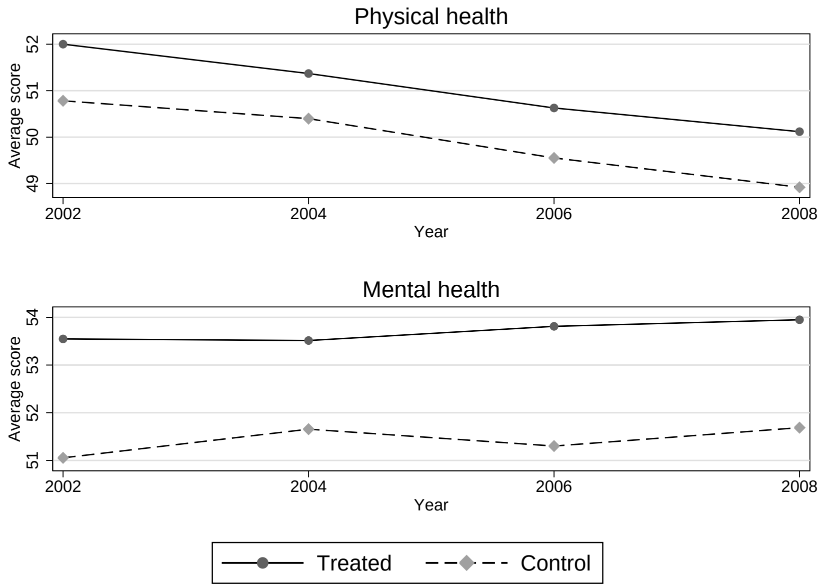


Figure A. 1: Pre-treatment trends in health outcomes

Appendix B Propensity-score model

Figure B. 1: Predicted probability of cultural-event attendance (at least once a month) by age.

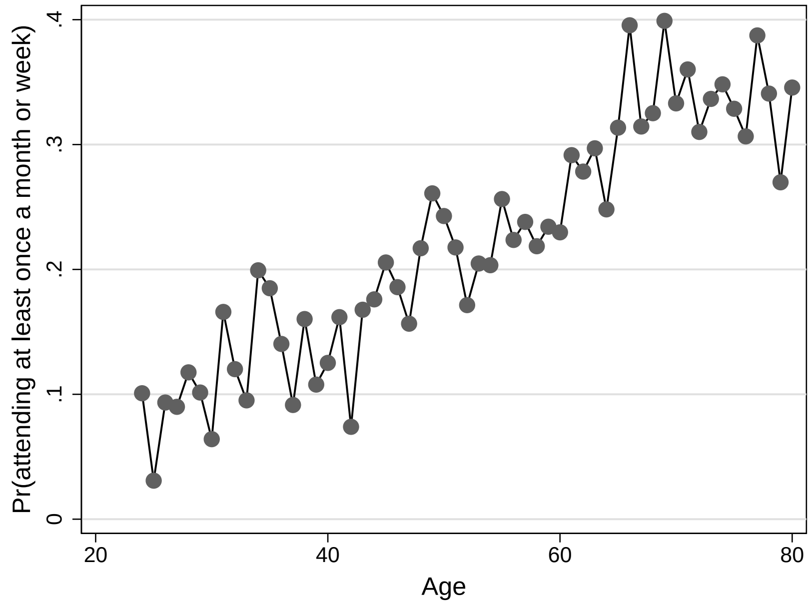
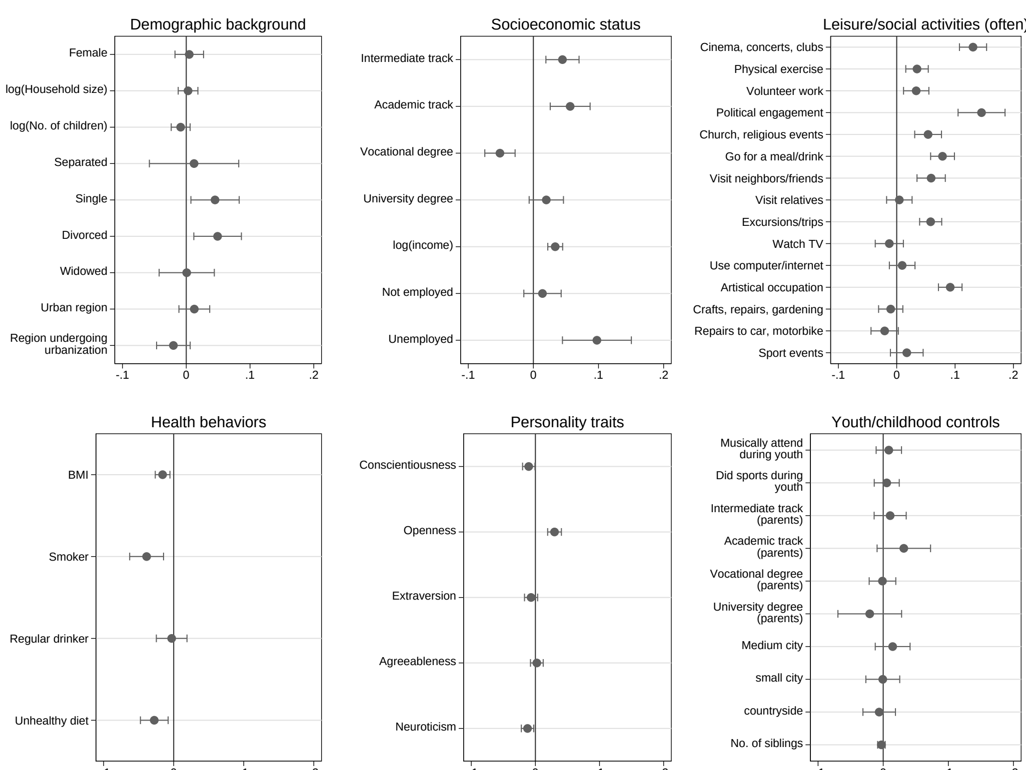


Figure B. 2: Determinants of cultural-event attendance (at least once a month), average marginal effects (90% CI)



Appendix C Robustness checks

C.1 Using different matching algorithms

The first set of robustness checks assesses the sensitivity of the empirical results to different specifications of the matching procedure. First, the matching models were reestimated 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 the causal-effect 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 similar matches, but uses fewer comparison units from the control group. However, as shown in panel A of Table (C.1), a lower bandwidth does not substantially alter the empirical estimates and the conclusions drawn in the previous section.

Panel B of Table (C.1) displays the ATTs from propensity-score kernel matching models using the Gaussian kernel and a bandwidth equal to 0.06. We do not find significant differences compared to the models using the Epanechnikov kernel, and the estimates remain fairly robust. This is in line with the observation that the choice of the kernel is relatively unimportant in practice, because any kernel will produce similar estimates of the probability density function (DiNardo and Tobias, 2001).

We also applied three variants of the nearest-neighbor matching algorithm (see, for example, Marcus, 2014; Morgan and Winship, 2007). The basic idea of nearest-neighbor matching is to find a match for each treated individual 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 control individuals. If only one nearest neighbor is used, the matched individual receives a weight equal to $W(i, j) = 1$. In case several nearest neighbors are used, the weight for the matched individuals equals $W(i, j) = 1/k_i$, where k_i is the maximum number of nearest neighbors assigned to each individual in the treatment group. The weight for unmatched control individuals is equal to 0. 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 treatment-effect estimates, but it could increase the bias due to the inclusion of poor matches.

We employed nearest-neighbor matching with replacement using 1, 5 and 10 nearest neighbors. Thus, an individual from the treatment group can be used multiple times for the calculation of the counterfactual of each treatment unit. Furthermore, to minimize the problem of using poor matches, we imposed a caliper equal to 0.001. 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.

Panel C to E of Table (C.1) show the results of the nearest-neighbor matching algorithm using 1, 5, and 10 nearest neighbors. The estimates are, at least qualitatively, comparable to the results using the kernel matching procedure. One notable exception is the non-significant treatment effect on mental health derived from the one-nearest-neighbor method (panel D) in model 2. This finding could, however, result from the loss of information and the increased variance when allowing only one match for each treatment case. As shown in panel D and E, the treatment-effect parameter for mental health in column 2 becomes more precise and significant. Hence, the non-significant estimate could just reflect the trade-off involved in choosing the number of nearest neighbors.

To summarize, the results presented in the previous section are rather insensitive to different specifications of the matching algorithm.

Table C.1: Robustness to matching algorithms

	Physical health			Mental health		
	(1)	(2)	(3)	(1)	(2)	(3)
A: Epanechnikov kernel, bw=0.03	0.465	0.501	-0.151	0.893**	0.975**	0.370
	(0.368)	(0.365)	(0.306)	(0.404)	(0.410)	(0.394)
N_{Treated}	932	932	932	932	934	934
N_{Controls}	3,283	3,283	3,283	3,283	3,283	3,283
$N_{\text{Off support}}$	21	21	21	21	19	19
Median bias (%)	1.6	1.7	1.7	1.6	1.5	1.5
B: Gaussian kernel, bw=0.06	0.445	0.457	-0.153	0.910**	0.993**	0.364
	(0.370)	(0.365)	(0.308)	(0.406)	(0.414)	(0.396)
N_{Treated}	932	932	932	932	934	934
N_{Controls}	3,283	3,283	3,283	3,283	3,283	3,283
$N_{\text{Off support}}$	21	21	21	21	19	19
Median bias (%)	1.6	1.6	1.6	1.6	1.5	1.5
C: Nearest neighbor, NN=1, cal=0.001	-0.070	0.540	0.060	0.630	0.628	0.163
	(0.466)	(0.452)	(0.404)	(0.489)	(0.467)	(0.490)
N_{Treated}	739	751	751	739	751	751
N_{Controls}	3,283	3,283	3,283	3,283	3,283	3,283
$N_{\text{Off support}}$	214	202	202	214	202	202
Median bias (%)	2.7	2.7	2.7	2.7	2.9	2.9
D: Nearest neighbor, NN=5, cal=0.001	0.174	0.463	0.136	0.732*	0.738*	0.267
	(0.397)	(0.393)	(0.345)	(0.397)	(0.399)	(0.411)
N_{Treated}	739	751	751	739	751	751
N_{Controls}	3,283	3,283	3,283	3,283	3,283	3,283
$N_{\text{Off support}}$	214	202	202	214	202	202
Median bias (%)	1.9	1.9	1.9	1.9	2.3	2.3
E: Nearest neighbor, NN=10, cal=0.001	0.195	0.513	0.136	0.783**	0.814**	0.267
	(0.389)	(0.386)	(0.345)	(0.390)	(0.394)	(0.411)
N_{Treated}	739	751	751	739	751	751
N_{Controls}	3,283	3,283	3,283	3,283	3,283	3,283
$N_{\text{Off support}}$	214	202	202	214	202	202
Median bias (%)	1.9	2.1	2.1	1.9	2.0	2.0

Robust standard errors in parentheses. * $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$. All models condition on demographic background, socioeconomic status, social and leisure-time activities, health behavior, personality traits, and childhood exposure to the arts. Column (1) displays results from the cross-sectional matching estimator using only the contemporaneous information on health outcomes in 2010. Column (2) shows estimates from the cross-sectional matching estimator including pre-treatment linear trends in health outcomes from 2002 to 2008 among the control variables. Column (3) represents the difference-in-difference matching estimator using the change in health outcomes between 2008 (pre-treatment) and 2010 (post-treatment) as outcome variables.

C.2 Including only significant covariates

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 the treatment as the outcome and successively add the covariates. A covariate is kept when it is significant at the 10 percent level. Table (C.2) displays the results from the matching analysis using only the subset of covariates which were significant in the propensity-score model. However, we were only able to reestimate the first (cross-sectional) and third (difference-in-difference) models because the pretreatment trends in health outcomes were not significant in the assignment models. Nevertheless, the estimates using only significant control variables are fairly similar and support the conclusions drawn in the previous section.

Table C.2: Robustness to the conditioning set: significant covariates only

	(1) Cross-sectional	(2) + pre-treatment health trends	(3) DiD
Physical health			
ATT	-0.046	-0.046	-0.094
	(0.382)	(0.382)	(0.341)
N_{Treated}	936	936	936
N_{Controls}	3,283	3,283	3,283
$N_{\text{Off support}}$	17	17	17
Median bias (%)	1.3	1.3	1.3
Mental health			
ATT	0.988**	0.988**	0.413
	(0.424)	(0.424)	(0.431)
N_{Treated}	936	936	936
N_{Controls}	3,283	3,283	3,283
$N_{\text{Off support}}$	17	17	17
Median bias (%)	1.3	1.3	1.3

[a] The pre-treatment outcome trends were not significant.

Robust standard errors in parentheses. * $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$. All models condition on demographic background, socioeconomic status, social and leisure-time activities, health behavior, personality traits, and childhood exposure to the arts.

C.3 Excluding potentially endogenous covariates

A further issue relates to the question whether some of the control variables are endogenous, that is whether they are influenced by the treatment variable. As argued by, for example, Rosenbaum (1984), the conditioning set should only include variables that are unaffected by the treatment or the anticipation of it. Lechner (2008) discusses the consequences of including such variables. He shows that the presence of control variables systematically influenced by the treatment can increase the bias of ATT estimates.

The control variables in our estimation sample are either time-invariant, measured prior to the treatment, or relate to conditions during childhood and youth. However, it is conceivable that some covariates can change due to the anticipation of the treatment in the following year. This particularly holds for the leisure-time and social activities among the control variables, which are measured the year before the treatment. For example, it could be possible that individuals already know that they will visit a classical concert or an art exhibition next year. Such highbrow events are often expensive and the individuals might, in the anticipation of next year's event, adjust their behavior in the current year. Hence, they probably spend less on leisure-time activities such as visiting sport events or holidays, and reduce their consumption of leisure goods to spend more on and increase their consumption of cultural goods next year. Hence, we dropped the leisure-time and social-capital variables from the conditioning set and reestimated the matching models. Thus, the models condition on factors that are assumed to be exogenous with respect to the treatment. These are the variables that are either time-invariant or cannot be easily adjusted to the anticipation of the treatment.[6] As indicated by Table (C.3), the estimates do not change substantially. Therefore, we are safe to conclude that a possible influence of future cultural participation on current leisure-time and social activities does not bias the treatment-effect estimates.

[6] 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 behavior.

Table C.3: Robustness to the conditioning set: excluding potentially endogenous covariates

	(1) Cross-sectional	(2) + pre-treatment health trends	(3) DiD
Physical health			
ATT	0.500	0.500	0.027
	(0.372)	(0.367)	(0.308)
N_{Treated}	942	942	942
N_{Controls}	3,283	3,283	3,283
$N_{\text{Off support}}$	11	11	11
Median bias (%)	1.1	1.1	1.1
Mental health			
ATT	0.900**	1.001**	0.412
	(0.370)	(0.370)	(0.376)
N_{Treated}	942	940	940
N_{Controls}	3,283	3,283	3,283
$N_{\text{Off support}}$	11	13	13
Median bias (%)	1.1	1.2	1.2
Outcome	Y_i	Y_i	ΔY_i
Pre-treatment outcome trend	No	Yes	Yes

Robust standard errors in parentheses. * $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$. All models condition on demographic background, socioeconomic status, health behavior, personality traits, and childhood exposure to the arts.

C.4 Bootstrapped standard errors

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 cluster-robust standard errors from the weighted regressions.

Another approach for variance estimation is bootstrapping, particularly if standard errors are difficult to compute analytically. It generally involves estimating the distribution of treatment effects in multiple samples randomly drawn with replacement from the observed sample. In this way, one is able to approximate the sampling distribution of the population mean (Brownstone and Valletta, 2001). Nevertheless, it is questionable whether bootstrap is valid in the case of matching. Abadie and Imbens (2008) show that the standard bootstrap procedure fails to provide asymptotically valid standard errors in the case of nearest-neighbor matching with a fixed number of matches. The same authors, however, speculate that bootstrapping is valid in the case of kernel matching which is asymptotically linear and with which the number of matches increases with sample size.

Therefore, inference primarily relies on robust standard errors from weighted linear 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 to calculate the ATTs 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 $\hat{\beta}_b^*$.

Table (C.4) compares the robust standard errors from weighted linear regressions with the bootstrapped standard errors. The latter are numerically comparable to the former but slightly more conservative. Thus, taking the uncertainty associated with the propensity-score estimation to some extent into account does not invalidate our results.

Table C.4: Comparison of robust standard errors with bootstrapped standard errors

	(1) Cross-sectional	(2) + pre-treatment health trends	(3) DiD
Physical health			
Robust s.e.	0.367	0.365	0.306
Bootstrapped s.e.	0.388	0.386	0.322
Mental health			
Robust s.e.	0.403**	0.424**	0.399
Bootstrapped s.e.	0.433**	0.437**	0.420

* $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$. Significance relates to ATT estimates. Robust standard errors are standard errors from weighted linear regressions clustered at the individual level. Bootstrapped standard error are based on 1,999 replications. All models condition on demographic background, socioeconomic status, social and leisure-time activities, health behavior, personality traits, and childhood exposure to the arts.

Appendix D Effect on 2012 outcomes

Table D.1: Matching estimates using 2012 outcomes

	(1) Cross-sectional	(2) + pre-treatment health trends	(3) DiD
Physical health			
ATT	0.700	0.716	0.448
	(0.440)	(0.439)	(0.375)
N_{Treated}	807	806	806
N_{Controls}	2,795	2,795	2,795
$N_{\text{Off support}}$	33	34	34
Median bias (%)	1.8	1.8	1.8
Mental health			
ATT	0.100	0.150	0.357
	(0.413)	(0.412)	(0.443)
N_{Treated}	807	811	811
N_{Controls}	2,795	2,795	2,795
$N_{\text{Off support}}$	33	29	29
Median bias (%)	1.8	1.8	1.8

Robust standard errors in parentheses. * $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$. All models condition on demographic background, socioeconomic status, social and leisure-time activities, health behavior, personality traits, and childhood exposure to the arts.

Appendix E Multivalued treatment effects

Table E.1: Matching results: pairwise comparisons

	Physical health			Mental health		
	(1) Cross-sectional	(2) + pre-treatment health trends	(3) DiD	(1) Cross-sectional	(2) + pre-treatment health trends	(3) DiD
A: T4 vs. T1	1.930	1.879	-0.767	1.214	1.553	4.747***
	(1.392)	(1.559)	(1.418)	(1.444)	(1.407)	(1.311)
N_{Treated}	38	35	35	38	37	37
N_{Controls}	829	829	829	829	829	829
$N_{\text{Off support}}$	79	82	82	79	80	80
Median bias (%)	13.2	11.0	11.0	13.2	14.6	14.6
B: T4 vs. T2	1.746**	1.983**	0.015	0.372	0.537	1.750**
	(0.751)	(0.734)	(0.533)	(0.709)	(0.693)	(0.696)
N_{Treated}	105	105	105	105	105	105
N_{Controls}	1,946	1,946	1,946	1,946	1,946	1,946
$N_{\text{Off support}}$	12	12	12	12	12	12
Median bias (%)	4.6	3.7	3.7	4.6	4.5	4.5
C: T4 vs. T3	0.323	0.531	0.423	-0.604	-0.608	0.638
	(0.714)	(0.721)	(0.600)	(0.611)	(0.604)	(0.604)
N_{Treated}	114	115	115	114	114	114
N_{Controls}	746	746	746	746	746	746
$N_{\text{Off support}}$	3	2	2	3	3	3
Median bias (%)	2.6	2.8	2.8	2.6	2.4	2.4
D: T3 vs. T1	-0.595	-0.559	-0.418	2.627***	2.599***	3.072***
	(0.588)	(0.584)	(0.434)	(0.735)	(0.724)	(0.598)
N_{Treated}	736	736	736	736	728	728
N_{Controls}	957	957	957	957	957	957
$N_{\text{Off support}}$	100	100	100	100	108	108
Median bias (%)	6.0	6.2	6.2	6.0	6.0	6.0
E: T3 vs. T2	0.406	0.401	-0.095	0.631*	0.697*	-0.037
	(0.378)	(0.371)	(0.318)	(0.368)	(0.369)	(0.380)
N_{Treated}	821	822	822	821	823	823
N_{Controls}	2,326	2,326	2,326	2,326	2,326	2,326
$N_{\text{Off support}}$	15	14	14	15	13	13
Median bias (%)	1.2	1.2	1.2	1.2	1.2	1.2
F: T2 vs. T1	0.609	0.635	-0.072	1.280***	1.194***	1.478***
	(0.449)	(0.448)	(0.346)	(0.483)	(0.481)	(0.461)
N_{Treated}	2,315	2,315	2,315	2,315	2,313	2,313
N_{Controls}	957	957	957	957	957	957
$N_{\text{Off support}}$	11	11	11	11	13	13
Median bias (%)	2.4	2.4	2.4	2.4	2.4	2.4

Robust standard errors in parentheses. * $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$. T1="never", T2="less often", T3="at least once a month", T4="at least once a week". All models condition on demographic background, socioeconomic status, social and leisure-time activities, health behavior, personality traits, and childhood exposure to the arts.