NBER WORKING PAPER SERIES

THE LONG-RUN EFFECTS OF SPORTS CLUB VOUCHERS FOR PRIMARY SCHOOL CHILDREN

Jan Marcus
Thomas Siedler
Nicolas R. Ziebarth

Working Paper 28819
http://www.nber.org/papers/w28819

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
May 2021

We would like to thank Sascha O. Becker, John Cawley, Jan Feld, Paola Guiliano, Helena Holmlund, Olena Izhak, Tobias Klein, Fabian Lange, Michael Lechner, Magne Mogstad, Frauke Peter, Jörn-Steffen Pischke, Imran Rasul, Jens Ruhose, Harald Tauchmann, Rudolf Winter-Ebmer, and Christoph Wunder for their helpful comments and suggestions. In particular, we are grateful to Hannah von Glan for the collection of the sports club data, and Dan Torge Dammann for his invaluable support in helping to design and implement the YOLO survey. We would also like to thank the participants of various research seminars and conferences for their helpful comments and suggestions. We would like to express our gratitude to Kristin Seidel, Regina Thoß, and Klaus Wallmann for their efforts in providing us with the health examination data and answering our questions. We are also grateful to Mr. Richter and Ms. Mai from the Landessportbund Sachsen, Mr. Kluger from the Saechsisches Staatsministerium fuer Kultus, Mr. Ender from the Kultusminsterium Sachsen, Ms. Riedel from the Sozialministerium Sachsen, and Ms. Frenschkowski from the Saechsisches Staatsministerium fuer Soziales und Verbraucherschutz for their willingness to meet with us and answer all our questions. Neither we nor our employers have any relevant or material financial interests that relate to the research described in this paper. We take full responsibility for any remaining errors in and shortcomings of the paper. The views expressed herein are those of the authors and do not necessarily reflect the views of the National Bureau of Economic Research.

The Long-Run Effects of Sports Club Vouchers for Primary School Children
Jan Marcus, Thomas Siedler, and Nicolas R. Ziebarth
NBER Working Paper No. 28819
May 2021
JEL No. H71,I12,I14,I18,I28,I38,Z28

ABSTRACT

Starting in 2009, the German state of Saxony distributed sports club membership vouchers among all 33,000 third graders in the state. The policy's objective was to encourage them to develop a long-term habit of exercising. In 2018, we carried out a large register-based survey among several cohorts in Saxony and two neighboring states. Our difference-indifferences estimations show that, even after a decade, awareness of the voucher program was significantly higher in the treatment group. We also find that youth received and redeemed the vouchers. However, we do not find significant short- or long-term effects on sports club membership, physical activity, overweightness, or motor skills.

Jan Marcus
University of Hamburg
jan.marcus@uni-hamburg.de

Thomas Siedler
University of Hamburg
Department of Economics
Von-Melle-Park 5
20146 Hamburg
and IZA
thomas.siedler@uni-hamburg.de

Nicolas R. Ziebarth
Department of Policy Analysis and Management
Cornell University
2218 MVR
Ithaca, NY 14853
and IZA
and also NBER
nrz2@cornell.edu

1 Introduction

According to the World Health Organization (2020), childhood obesity is "one of the most serious public health challenges of the 21st century." To fight this phenomenon, it calls for coordinated and comprehensive policy action such as promoting healthy diets, taxing unhealthy food, and incorporating physical activity into the daily routines of children (Frieden et al., 2010; World Health Organization, 2017). However, we know very little about which policies actually work.[1] In particular, while there is consensus that a lack of physical activity is a major driver of children's unhealthy body weight (Prentice-Dunn and Prentice-Dunn, 2012), it remains largely unknown how to induce children to exercise more. This paper comprehensively evaluates a $5 million sports club voucher policy with the objective of nudging primary school children into exercising more and adopting exercise as a long-term habit.

Specifically, we use survey and administrative data to empirically evaluate the causal effects of this voucher policy experiment in the German state of Saxony. The policy's objective was to increase physical activity through the distribution of sports club vouchers and an information campaign among all third graders in Saxony in 2009, 2010, and 2011. The sports club vouchers were worth about $50-60 and provided free membership for up to one year. The campaign slogan was "Come to the Sports Club" (*KOMM! in den Sportverein*). The idea was to encourage children (and their parents) to test regular use of a sports club for free, with a view to them becoming active members and adopting a healthier lifestyle in the short- and long-run. One of the initiative's target groups were children from disadvantaged, economically deprived households who could not afford sports club membership fees (German Olympic Sports Confederation, 2011). In Germany, children from households in the lowest income quintile are about 50 percent less likely to be physically active in their leisure time than children from households in the highest quintile (Graf and Cecchini, 2019).

To evaluate the voucher program's effectiveness and its impact on awareness, membership take-up, physical activity, body weight, and health, we rely on two unique data sources, a

[1] Cawley et al. (2007) and Cawley et al. (2013) are exceptions exploiting variation in state-level mandated minimum physical education (PE) class time. They find lower obesity rates among boys in fifth grade (Cawley et al., 2013) and an impact on actual PE instruction time in high school, but no significant effect on high school students' body weight (Cawley et al., 2007). Otherwise, the rich economics obesity literature focuses on adult obesity and studies general driving forces such as technological change (Cutler et al., 2003; Lakdawalla and Philipson, 2009) along with specific drivers such as the availability of (fast food) restaurants (Currie et al., 2010; Dunn, 2010; Anderson and Matsa, 2011), consumption of soda (Fletcher et al., 2010), increases in portion sizes (Jeitschko and Pecchenino, 2006), increases in gluttony (Griffith et al., 2016), declining gas prices (Courtemanche, 2011), increasing cigarette taxes (Courtemanche, 2009), changes in food prices (Dubois et al., 2014; Grossman et al., 2014; Strulik, 2014; Courtemanche et al., 2015; Dragone and Ziebarth, 2017), and cash transfers (Akee et al., 2013).

register-based survey and administrative data from school health examinations. For the survey, we first contacted registry offices (*Einwohnermeldeämter*) in the German states of Saxony, Brandenburg, and Thuringia and obtained 80 percent random samples of residential addresses of treatment and control cohorts. In 2018, we then contacted these households by regular mail with an invitation to participate in an (incentivized) online survey, the *Youth Leisure Online Survey* (YOLO), which we designed for the purpose of this study. Around 16,000 youth completed the survey. We use registry data to compare characteristics of survey participants and non-participants and show that survey participation was not affected by the voucher program. Moreover, we show that the distribution of the socio-demographics in YOLO is similar to that of the representative German Socio-Economic Panel (SOEP). For the administrative data set, we exploit that state employed physicians are legally required to physically examine all children in school and obtained complete examination data for several cohorts in one county of Saxony. The data include objective measures on overweightness and obesity as well as emotional and motor skill disorders of about 7,000 sixth graders, some of which were and some of which were not treated.

Using difference-in-differences models and comparisons across affected and unaffected cohorts, our findings show that the "Come to the Sports Club" (C2SC) campaign has been effective in increasing long-term awareness about the program, especially among the first treated cohort. Even seven to nine years after the program started, significantly more treated children in Saxony recall having received and redeemed the vouchers, relative to older cohorts and cohorts in neighboring states. However, despite higher awareness and utilization, we find no significant short- or long-term effects on membership rates, physical activity, and overweightness among previously inactive students. Conversely, we find strong evidence that the vouchers were a windfall gain for parents of existing members as they were the ones who primarily redeemed the vouchers. Consistently, we do not find that the voucher program changed self-reported health, health behaviors, objectively measured obesity rates, and diagnosed disorders either in the short or long-run. We discuss several potential explanations for the ineffectiveness of the program and perform a battery of robustness tests including synthetic control group methods and placebo checks.

There are several arguments to suggest that this unique sports club voucher program was not poorly designed, but rather a well-crafted and implemented nudging policy. First, it focused on primary school children. Childhood is a crucial age for habit formation and children

are more receptive to interventions than adults (Just and Price, 2013; Belot et al., 2016; Loewenstein et al., 2016). Second, as the voucher provided free sports club membership for one year, relative to existing interventions in the literature, it is quite a long-run intervention. This is a crucial and inherent element of C2SC as changing habits takes time (cf. Royer et al., 2015; Carrera et al., 2018, 2020). Third, German sports clubs usually focus on team sports. Hence, because of peer pressure from other team members and coaches, and out of a feeling of responsibility toward their team, children might have stayed active members for an extended time period (Babcock et al., 2015). Fourth, German sports clubs typically offer sports activities for all age groups; children can therefore remain active members for many years (Breuer et al., 2015). Fifth, being a member of a sports club implies regular, weekly practice. For example, among German fifth graders who are sports club members, 36 percent report practicing several times a week (Züchner and Arnoldt, 2012). Sixth, the voucher was not a cash transfer but an in-kind transfer and could not be used for any other purpose than sports club membership in the given time frame (see, e.g., Currie and Gahvari, 2008).

In addition to contributing to the very sparse literature on how to increase school children's physical activity levels, this research contributes to the broader economics literature on nudges for adults to adopt a healthier lifestyle. Studies show that adults can be incentivized to go to the gym more often, but the effects are short-lived and people seem to find it difficult to change their habits permanently (DellaVigna and Malmendier, 2006; Royer et al., 2015; Carrera et al., 2018, 2020). Exceptions are Reichert (2015) and Augurzky et al. (2018) who find long(er)-lasting effects on weight loss and health behavior in a randomized controlled trial among obese health plan enrollees in Germany. In one of the few studies that focuses on children, Angelucci et al. (2020) study peer effects in health behaviors in a field experiment among K-8 school children in Chicago. Their findings demonstrate the existence of health behavior spillover effects, but also that making incentives public can backfire and crowd-out positive peer effects.[2] Because we evaluate a policy that explicitly targeted primary school children with the intention of changing their health behavior in the long-run, our research also contributes to research on early childhood interventions which have been shown to have long-lasting effects (see, e.g., Kesternich et al. (2015) and Felfe and Lalive (2018) for two papers with a focus on children in Germany and similar data).

[2] In another study, Prina and Royer (2014) find that body weight report cards increase parental knowledge about their children's body weight without having any impact on parental behavior or children's weight.

The next section explains the extramural sports club setting in Germany and describes the voucher policy in detail. Section 3 discusses the data and key variables, and Section 4 explains the empirical models. Section 5 presents our findings based on YOLO, and Section 6 presents robustness checks. Section 7 discusses additional evidence based on official health examination data, and Section 8 concludes.

2 Non-Profit Sports Clubs and the Voucher Initiative

2.1 Germany's Extramural Sports Club Setting

Germany has a long tradition of extramural sports clubs. In 2018, there was a total of 89,121 sports clubs in Germany distributed over 11,000 municipalities (Breuer and Feiler, 2015).[3] Sports clubs are the main providers of opportunities for organized sport and cover all ages; for instance, about 4.2 million members are between the ages of 7 and 14, relative to roughly 6 million Germans in that age group (German Olympic Sports Confederation, 2017).

Unlike in the United States, German extramural sports clubs are not associated with primary or secondary schools but operate as independent, voluntary, non-profit amateur organizations.[4] Coaches are typically former or current amateur athletes. In general, the clubs charge low membership fees and admit anyone who applies to be a member. Bigger clubs in larger cities may participate in professional or semi-professional sports leagues.

Most members join extramural sports clubs as children or youths between the age of 5 and 15. If a young person is serious about a particular sport, that is, if they decide they want to acquire, develop, and hone the skills the sport involves, they will typically practice several times a week, join a team in their age group, and participate in amateur competitions. This aspect of the extramural sports clubs closely resembles the sort of activities young people in the US pursue when they join school-based sports teams. As in the US, German youth compete in matches that take place in their hometown, and they travel to compete in matches hosted by clubs of other towns in their state of residence. The competitions and the associated travel

[3] Breuer and Feiler (2015) and Breuer et al. (2015) provide a summary of the history and organization of extramural sports clubs in Germany.

[4] According to Breuer et al. (2015), sports clubs in Germany can be characterized by several constitutive and economic features. The constitutive characteristics are: membership is voluntary and members can freely decide when to enter and when to leave the clubs, sports clubs are autonomous, focus on the interests of their members, have democratic decision-making structures, and rely on volunteers. The main economic features are their non-profit orientation (clubs are allowed to make a profit, but they are not allowed to pay out surpluses to members), autonomous revenues, and the principle of solidarity.

are an integral part of the experience young people gain when they participate in extramural sport.

2.2 The Voucher Initiative

On June 18, 2008, during his State of the Union Address, the newly sworn-in prime minister of the east German state of Saxony, Stanislaw Tillich (Christian Democratic Union, CDU), announced a new policy initiative. The main goal of this "Come to the Sports Club" (C2SC; *KOMM! in den Sportverein*) initiative was to induce primary school children to join an extramural sports club. By joining a sports club, children would not only exercise regularly, be healthier, and more self-confident, but also meet new friends and acquire social skills to cope better with everyday life and become "good citizens". The C2SC initiative was jointly developed by the Ministry for Education and Cultural Affairs (*Kultusministerium*) and the Saxony State Sports Association (*Landessportbund Sachsen*).

To not discriminate against anyone but ensure that low-income families were in a position to afford the membership fees, the idea was to distribute membership vouchers among all 33,000 third graders in Saxony at the end of January 2009. Figure A1 (Appendix) shows an example of the voucher. The vouchers were handed out by primary school class teachers and had the official school stamp to prevent illegal copies from being made. They were distributed together with a "starter kit", which included a T-shirt with the logo of the initiative as well as an information letter for the parents describing the basic idea of the initiative and that the voucher could be redeemed in any of the state-approved sports clubs until March 31, 2009. The letter also explained that a second voucher would be distributed at the beginning of the fourth grade in August 2009 (the idea was to let children experiment with several sports clubs and disciplines). Both vouchers were worth €30 ($33), so children of this cohort received vouchers worth €60 ($66) in total. The voucher's value was designed to cover membership fees in the majority of sports clubs: Breuer and Feiler (2015) report that the annual median membership fee was €30 for children in Germany in 2013. Moreover, the letter informed parents that sports club membership included insurance for sports injuries and referred to brochures with complete lists, addresses, and all disciplines offered for children by all local sports clubs in Saxony. The brochures were distributed in all primary schools as part of the information package.

The letter also referred to a website and a contact person (including an e-mail address and a telephone number).

The initiative was repeated for the following two years. That is, three cohorts were treated, namely those who were third graders in Saxony in school years 2008/2009, 2009/2010, and 2010/2011. While in the first year of the initiative, two vouchers worth €30 each were distributed, in the second and third year, only one voucher worth €50 was distributed in late January 2010 and 2011, respectively. The reason for switching from two vouchers to one voucher was to reduce the administrative burden for the sports clubs and to align the funding with the clubs' fiscal year (calendar year). The deadline for redeeming the voucher in the second and third round was again March 31.[5]

Based on the number of redeemed vouchers, each sports club was reimbursed for the "lost" membership fees. At the beginning of the third round, in January 2011, the initiative announced that about 20,000 vouchers (out of a total of about 66,000 eligible third graders) had been redeemed. The budget for the entire C2SC initiative was €4.5 million over three years (German Olympic Sports Confederation, 2011).

In 2012, C2SC was restructured and third graders no longer received sports club vouchers. Instead, the vouchers were distributed to first graders.[6] In 2013, policymakers decided to completely abolish the voucher system, primarily because of the high administrative burden for distributing the reimbursements. The C2SC program still exists today, however. It has been broadened and now also focuses on adolescents and people over the age of 50. Instead of distributing vouchers, regional coordinators were hired to build "regional networks to foster physical activity among the population" (Kreissportbund Landkreis Leipzig, 2019; Landessportbund Sachsen, 2019).

[5] In 2011, the German federal government launched the Educational Package (*Bildungs- und Teilhabepaket*), which among other things, covers membership fees and equipment costs for sports clubs. However, this program was only directed at low-income welfare recipients and came into effect on April 1, 2011, that is, after the deadline for redeeming the voucher from the last round of the initiative. Our results are robust to only including the first two rounds of C2SC (see Section 6).

[6] Note that we do not analyze the C2SC effects on first graders as the affected students are mostly below the age of 14 at the time of our survey. In Germany, children below the age of 14 may not be surveyed without parental consent.

3 The Youth Leisure Online Survey (YOLO)

This paper relies mainly on a large register-based online survey, which we specifically designed and carried out to evaluate the effectiveness of the "Come to the Sports Club" (C2SC) initiative. This main data set, the **Yo**uth **L**eisure **O**nline Survey (YOLO), uses a two-stage sampling strategy, which we describe in detail in the Appendix and Figure A3. By law, Germany requires residents to register with the registry offices (*Einwohnermeldeämter*) of their municipality; addresses of households can be obtained for research purposes.[7] Hence, we had to randomly sample at the level of the registry offices and request addresses for a sample of households with youth of our target group (registry offices know the demographics of each household member). We could then contact these households via regular mail and invite them to participate in an online survey.

Hence, in the first stage, we randomly sampled 121 registry offices in Saxony and the neighboring states of Brandenburg and Thuringia (with sampling probabilities proportional to population size).[8] Figure A2 (Appendix) illustrates the geographic location of the registry offices along with the population that they cover in the three federal states.

In the second stage, we contacted all 94 registry offices who responded in the first stage, and requested an 80 percent random sample of the target population. This target population consists of individuals born between July 1997 and June 2002. The first cohort that received the C2SC voucher was born between July 1999 and June 2000. These individuals typically entered school in 2006 and the third grade in the school year 2008/2009. Table A1 (Appendix) lists the relationship between birth cohorts and school cohorts. We aimed to survey two pre-voucher cohorts (born July 1997 to June 1999) and three treated voucher cohorts (born July 1999 to June 2002) in Saxony. Moreover, we surveyed the same five cohorts in the neighboring states of Brandenburg and Thuringia.[9]

[7] While most municipalities have their own registry office, some registry offices are responsible for more than one municipality. For ease of understanding, we will use the terms registry office and municipality interchangeably throughout the paper.

[8] We chose these three states because they all use the same registry software which facilitated the execution of the survey and allowed us to provide the registries with instructions on how to randomly draw the subsamples. The neighboring state of Saxony-Anhalt uses a different software and is therefore not included. 94 of the 121 registry offices contacted responded. The response rate of the registry offices was similar in Saxony (77.5 percent) and the control states (77.8 percent) and did not differ between urban and rural regions.

[9] Note that we also surveyed the cohort born between July 2002 and June 2003, that is, third graders in school year 2011/12. In our main analysis, we discard this cohort for two reasons. First, the Educational Package (*Bildungs- und Teilhabepaket*) that covered sports club membership fees for welfare recipients came into effect on April 1, 2011 and therefore affected this cohort (see Section 2.2). Second, this is the first cohort that did not receive vouchers and, therefore, disappointment effects could arise. However, in Section 6 we show that our findings are robust to including this post-treatment cohort.

Next, we mailed one official invitation letter to each of the 155,527 adolescents sampled in the second stage. Figure A4 (Appendix) shows the original invitation letter for an online survey about youth leisure time activities. To increase response rates, we offered a lottery ticket for participation.[10] This letter provided a unique access code for the online survey for both, children *and* their parents.[11] This unique access code also allowed us to match children with their parents. Respondents completed the surveys between March and July, 2018. It took respondents an average of 34 minutes to complete the survey (see Figure A5, Appendix). The final YOLO response rate was 12.7 percent. Below, we investigate and discuss selective survey participation. For this purpose, we also make use of the representative German Socio-Economic Panel Study (SOEP).

In Section 7, in addition to YOLO, we use official student health examination data carried out by state-employed physicians. We use this second data set from one county in Saxony to examine objectively measured body weights, heights, as well as motor skill and emotional disorders.

3.1 Main Outcome Variables

In total, we generate six main outcome variables: three measure policy awareness and voucher utilization (*program known, voucher received, voucher redeemed*), and three measure sports club membership, physical activity, and overweightness (*member of sports club, weekly hours of sport, overweight*). We elicited all six outcomes at the time of the survey. However, the first three are of retrospective nature, whereas the other three are of contemporaneous nature.

Note that the invitation letter *did not* specifically mention the C2SC campaign (see Figure A4, Appendix). It only stated that the survey would be about leisure time behavior among young people. Further, so as not to frame participants, the questions regarding the first three awareness and utilization measures (*program known, voucher received, voucher redeemed*) appear only at the very end of the survey.

3.2 Sample Selection

Our main sample consists of young people who attended third grade in primary school in Saxony, Thuringia, or Brandenburg during the school years from 2006/07 to 2010/11. Hence,

[10] The lottery prize included two iPads, worth €500 each, and ten Amazon vouchers, worth €20 each.
[11] We invited one parent (mother or father) to participate in the survey.

we work with two pre-treatment and three treatment cohorts. YOLO explicitly asks when children were born, in which state they attended each primary school grade, whether they started first grade at the age of six, and whether they had to repeat a grade. This allows us to precisely assign respondents to the treatment group. For example, we disregard individuals who attended third grade abroad or in other federal states. Further, we excluded observations with missing values on one of our six outcome variables. Our final sample consists of 13,334 unique youth observations.

3.3 Descriptives

Table 1 shows the summary statistics of the YOLO sample. The average age of the young people surveyed was 17.5 years. 57 percent of all respondents were female. About half were attending an academic track school at the time of the survey in 2018.[12] Instead of asking respondents about their parents' income or education, to proxy for socio-economic status, we asked whether they had a newspaper and/or art at home. Of all the respondents, 58 percent had a newspaper at home and 73 percent had art at home.

38 percent of all respondents belong to the treatment group, that is, they were third graders in Saxony between school years 2006/07 and 2010/11. Table 1 shows that 19 percent of all respondents (including control cohorts) had heard about C2SC (*program known*). Almost ten percent recall having received the voucher (*voucher received*) and six percent had redeemed the voucher (*voucher redeemed*). Around 42 percent of youth were active extramural sports club members at the time of the survey (*member of sports club*). On average, respondents exercised 4.6 hours per week (*weekly hours of sport*) and 16 percent were *overweight*.

Figure B1 (Appendix) plots the sports disciplines for which the vouchers were redeemed. Not surprisingly, by far the most popular discipline was soccer (20 percent), followed by martial arts (12 percent), swimming (12 percent), handball (7 percent), athletics (6 percent), and gymnastics (5 percent).

[Table 1 about here]

Table A2 (Appendix) lists the normalized difference for key variables between treatment and control groups. According to Imbens and Wooldridge (2009), a normalized difference of

[12] The states considered here track students after four years of mixed primary school.

more than 0.25 indicates strong imbalances. As seen, almost all variables are balanced across the treated and control groups, with the majority of the normalized differences smaller than 0.10. For example, the average age at the time of the survey was 17.6 (treated) vs. 17.4 (control); about 57 percent of respondents were female in both groups, and 20.5 percent vs. 21.8 percent had ever smoked a cigarette.[13] Moreover, the sport-related outcome variables are very balanced, whereas program-related outcome variables naturally differ.

3.4 Data Quality

To check the accuracy of the information provided by the YOLO respondents and to investigate possible selective survey response, we use two secondary data sources: the registry data provided by the registry offices and the German Socio-Economic Panel Study (Goebel et al., 2019). First, we have a set of overlapping variables for YOLO participants, as the registry offices provide administrative information which we also surveyed in YOLO. YOLO participants were not aware that we had the registry information. Table A3 (Appendix) shows that 99 percent of all YOLO participants correctly reported their gender, nationality, and year of birth.[14]

Second, based on the registry data, Table A4 (Appendix) compares the characteristics of YOLO participants and non-participants. While YOLO has a slight overrepresentation of women, Germans, and younger individuals, all normalized differences are below 0.25, suggesting that participants and non-participants do not differ strongly with respect to these characteristics. Additionally, in Section 6, we show that these slight overrepresentations are unrelated to C2SC and that the results are robust to weighting the observations based on their probability of participating in the survey.

Third, we use the SOEP as a reference data set to compare YOLO and SOEP participants (Siedler et al., 2009). We can directly compare a wide range of background information between YOLO and the SOEP as the wording of several YOLO questions is identical to the wording in the SOEP (including socio-demographic variables, leisure time activities, personality traits, and attitudes). Specifically, we compare SOEP and YOLO respondents who were born between July

[13] One of the very few covariates with a major imbalance is the share of youth living in a city (76 percent in the treatment group vs. 43 percent in the control group). The reason for this is that Saxony has two cities with more than half a million residents (Dresden and Leipzig), whereas Brandenburg and Thuringia do not.

[14] Four registry offices provided only the respondent's gender and address but not the date of birth. Therefore, the number of observations is slightly smaller for the birth variables in Table A3 (Appendix). In the robustness section, we show that the results are robust to excluding individuals with non-matching registry information and individuals who spent little time answering the survey.

1997 and July 2000 and who live in Saxony, Brandenburg, or Thuringia.[15] Table A5 (Appendix) compares an extensive list of covariates, such as demographics, sports and other leisure time activities, volunteering, personality traits, and attitudes. Again, most variables are very balanced and do not provide much evidence for strong selection into YOLO.

4 Empirical Approach

We estimate the C2SC effects using difference-in-differences (DD) models. We begin with basic DD models of the form:

$$Y_{ics} = \alpha_0 + \alpha_1 \cdot Saxony_s + \alpha_2 \cdot Post_c + \beta \cdot (Saxony \cdot Post)_{cs} + \varepsilon_{ics}, \quad (1)$$

where Y_{ics} denotes the outcome of individual i in cohort c in state s. The dummy variable $Saxony_s$ is one if a respondent was a third grader in Saxony, and zero if he or she was a third grader in Thuringia or Brandenburg. Recall that about 33 thousand third graders of the treatment cohorts were treated and received the voucher. $Post_c$ is a dummy variable, which is one if a respondent was a third grader in the school year 2008/09, 2009/10 or 2010/11, and zero if he or she was a third grader in the school year 2006/07 or 2007/08. The main variable of interest is the interaction of these two dummies ($Saxony \cdot Post$). It is one if a respondent was eligible to receive and redeem sports club vouchers, and zero for all other respondents. ε_{ics} is the individual error term. To take the nature of our sampling strategy into account, we cluster our standard errors at the level of the municipalities and, hence, allow for arbitrary correlation of these error terms across municipalities, the primary sampling units of our survey (Abadie et al., 2017).[16]

In our second specification, we replace the $Saxony_s$-dummy with state fixed effects (κ_s) and the $Post_c$-dummy with cohort fixed effects (γ_c). This twoway-fixed effects specification considers general differences in the outcomes between states as well as general changes in the outcomes over time. The estimation equation is then:

[15] We rely on the SOEP youth questionnaire, which surveys individuals in the year in which they turn 17. This fixed age is an important difference between SOEP and YOLO participants. While the average age of YOLO respondents is comparable, due to the sampling design, YOLO respondents are between the ages of 14 and 20 at the time of the survey (see also Table 1). As many older YOLO participants are no longer in school, we refrain from comparing school-related variables.

[16] In Section 6, we present and discuss several alternative methods of inference.

$$Y_{ics} = \beta \cdot (Saxony \cdot Post)_{cs} + \gamma_c + \kappa_s + \varepsilon_{ics}. \tag{2}$$

In our third and preferred specification, we augment equation (2) with municipality fixed effects.

The main identifying assumption of our DD models is the common time trend assumption. This means that, in the absence of C2SC, the outcomes of the treatment and control group would have followed the same time trend. Below, we provide support for this assumption graphically and by running placebo regressions using unaffected cohorts and unaffected states.

5 Results

This section first presents graphical and regression-based evidence of C2SC's effectiveness (Section 5.1) and then examines effect heterogeneity between subgroups (Section 5.2). Next, we investigate the program's short-run effects using retrospective information about sports club membership throughout the participant's childhood (Section 5.3). Section 5.4 presents suggestive evidence for potential mechanisms.

5.1 Main Results

We begin by plotting unadjusted mean outcomes separately for the treatment state (Saxony) and the control states (Brandenburg and Thuringia) in Figure 1. We plot these means separately by the school year during which YOLO respondents attended third grade. The three subgraphs in the left column of Figure 1 illustrate whether treated cohorts in Saxony have a higher awareness of the program (1a), were more likely to have received the voucher (1b), and were more likely to have redeemed the voucher (1c).

As seen, we observe very flat and non-trending lines over the entire time period for respondents in the control states for all three outcomes (dashed lines).[17] In contrast, respondents who were third graders in Saxony at the time of the policy show a substantially larger program

[17] In Figure 1a, some individuals in Saxony's pre-treatment cohorts report to have known the program, and in Figures 1b and 1c a small fraction of individuals in Saxony's pre-treatment cohort claims that they received and used the voucher. There are several potential explanations for this phenomenon. For instance, it could be that the children got hold of the voucher although they were not eligible (e.g., via siblings or friends or a teacher handing out the voucher to the wrong class), or that individuals remember incorrectly whether they received the voucher (recall bias).

awareness, particularly those who were treated in the first year of the policy. While awareness clearly decreases for the two following cohorts from about 50 percent to 40 percent and 30 percent, respectively, it still remains higher than in control cohorts and control states. Interestingly and reassuringly, we observe the exact same pattern for the outcomes *voucher received* and *voucher redeemed* in Figures 1b and c: There are substantial spikes in the first year of the voucher initiative, and subsequent linear decreases in years two and three. Note that the decreases in program awareness, treatment, and utilization works against a possible recall bias (which would increase in the years elapsed since then). The dynamically decreasing treatment effects are more likely to be a function of the very active information campaign and promotion in the first C2SC year. Also, C2SC's structure changed from disbursing two smaller vouchers (which were valid for six months) in January 2009 and August 2009, to one larger voucher (which was valid for 12 months) in January 2010 and 2011 (see Section 2.2 for details).

[Insert Figure 1 about here]

The three subgraphs in the right column of Figure 1—Figures 1d, e, and f—show the three outcomes *membership in sports club today*, *weekly hours of sports*, and *overweight*. Here, no treatment effect is visually detectable. The two lines follow almost identical trends and levels throughout the entire time periods. Note that, while the first three awareness and utilization measures are elicited retrospectively, the three sports club and activity measures are elicited *contemporaneously* (and thus do not suffer from any recall bias).

Next, we turn to our main regression results obtained with the DD models. We start with the simplest of all specifications in column (1) of Table 2, where we only include a binary treatment group indicator (which is one for respondents who attended third grade in Saxony), a binary post-reform indicator (which is one for school years 2009/10 and after), and its interactions (equation (1)). Column (2) adds a full set of cohort and state fixed effects (equation (2)), and column (3) additionally includes municipality fixed effects. Standard errors are clustered at the municipality level.

[Insert Table 2 about here]

The findings from the 18 DD models in Table 2 allow us to conclude the following: First, the regression results are entirely in line with the visual evidence above: Awareness about the program is on average 27 percentage points higher among those who attended third grade in

Saxony during the voucher years. In line with the nonparametric evidence in Figure 1, treated cohorts also have a 20 percentage point higher probability of reporting that they received the voucher, and a 12 percentage point higher probability of reporting that they redeemed the voucher. In Panel A, all nine coefficient estimates are highly significant at the one percent level. Moreover, and again in line with the graphical evidence, the DD models in Panel B show that the program was neither effective in increasing sports club membership rates or physical activity, nor in reducing overweight rates *in the long-run* (seven to nine years later). Second, the coefficients are very robust to the inclusion of additional time and region fixed effects, suggesting the absence of confounding time trends or spatial factors.

As we surveyed all respondents in 2018, some of our dependent variables may suffer from recall bias due to the retrospective nature of our survey questions. While such recall bias is unavoidable, it is important for the consistency of our estimates that no systematic, treatment-related, recall bias exists. Note that our main dependent variables fall into two categories: (i) The outcomes *voucher received* and *voucher redeemed* are likely to suffer from recall bias, as respondents who received and redeemed the vouchers are more likely not to recall that this was the case. This almost exclusively affects the treatment group as the control group did not receive any vouchers. Hence, the take-up estimates are likely to be downward biased and yield a lower bound.[18] While it could be argued that the recall bias is smaller when using parental information, we obtain similar point estimates when using parents' rather than their children's responses (see Section 6); (ii) When estimating the long-run effects of the C2SC voucher policy on *contemporaneous* sports club membership rates, physical activity, and overweightness, the estimates cannot be affected by recall bias as they are not retrospective.

5.2 Effect Heterogeneity

Now we investigate effect heterogeneity to better understand the underlying driving forces of the treatment effects and to examine whether treatment effects differ between specific sub-groups. Technically, we run a version of equation (2) with municipality fixed effects, to which

[18] For the outcome *voucher received* we would expect a point estimate close to one if the program was perfectly administered and if there was no recall bias. Regarding the outcome *voucher redeemed*, official numbers suggest that about 30 percent of eligible students redeemed the voucher in the first two years of the program: 20,000 vouchers were redeemed and about 66,000 third graders were eligible (German Olympic Sports Confederation, 2011).

we add interactions of all variables with the stratifying variable of interest:

$$Y_{igcs} = \beta_1 \cdot Voucher_{cs} + \beta_2 \cdot Voucher_{cs} \cdot Group_{igcs} + \gamma_{cg} + \kappa_{sg} + \lambda \cdot Group_{igcs} + \varepsilon_{igcs}, \quad (3)$$

where Y_{igcs} denotes the outcome of individual i in cohort c in state s and group g. γ_{cg} and κ_{sg} are cohort-group and state-group fixed effects. *Group* is an indicator that is one if an individual belongs to a specific group (e.g., females) and zero if not (e.g., males). Figure 2 graphically plots the $\hat{\beta}_2$-coefficients, i.e., the difference in the treatment effects between groups for six different stratifying variables and all six outcomes along with the 95 percent confidence bands. Additionally, Table B1 (Appendix) shows the DD regression coefficients for all subgroups.[19]

[Insert Figure 2 about here]

As above, we begin with the three measures for program awareness, take-up, and utilization in the left column of Figure 2. A clear picture emerges. While we do not find much evidence that the effects differ by gender or urban/rural regions, all effect sizes are significantly larger for children from higher socio-demographic backgrounds (i.e., children whose parents have art and/or a newspaper at home). They are also larger for youth attending an academic track school and those who were already sports club members before the C2SC campaign started. Although effect sizes differ, in line with our main findings, Table B1 (Appendix) shows positive and significant effects for program awareness, take-up, and utilization for every subgroup. They strongly reinforce the narrative that it was primarily children from advantaged parental backgrounds and existing sports club members who redeemed the voucher, but not the main target group of disadvantaged children from economically deprived households.

The right column of Figure 2 shows effect heterogeneity for the three contemporaneous long-term measures: sports club membership, physical activity, and overweightness. They confirm a lack of significant long-term effects of the program. Table B1 (Appendix) shows no single significant treatment effect for these long-term outcomes for any of the subgroups.

[Insert Figure 3 about here]

As another heterogeneity test, we examine effect heterogeneity by cohorts and plot the results in Figure 3.[20] The results in Figure 3 reflect and reinforce the nonparametric evidence from

[19] Technically, Table B1 (Appendix) is based on separate DD regressions for each group, which yields estimates for β_1 and $\beta_1 + \beta_2$.

[20] For this purpose, we add cohort-specific treatment indicators to equation (2) with municipality fixed effects:

Figure 1: The highly significant effects on awareness, take-up, and utilization are strongest for the first affected cohort and then decline substantially for the following two treated cohorts. Again, we find no evidence for significant long-term effects on membership rates, physical exercise, or being overweight for any of the treated cohorts.

5.3 Short-Run Effects on Sports Club Membership Rates

While the previous sections demonstrated an absence of long-run effects on sports club membership rates, this section examines potential short-run effects. To this end, we make use of the retrospectively reported membership information by child age from 5 to 12.

[Insert Figure 4 about here]

Figure 4a shows the nonparametric development of membership rates by child age 5 to 12. Here we focus only on YOLO respondents who went to primary school in Saxony and show the results separately for treated and non-treated cohorts. First, we see monotonically increasing and concavely shaped membership rates as a function of child age. Between ages 5 and 8, the probability of being a sports club member roughly doubles from 25 to 50 percent. It then flattens substantially between ages 9 and 12 but continues to increase. Moreover, both lines are almost identical and clearly follow a common time trend—both before and after age 9 (the age when children are third graders). We observe the same concave-shaped function in Figure 4b, which focuses solely on respondents who were third graders at the time of the C2SC initiative and compare the treatment state of Saxony to the control states of Brandenburg and Thuringia. The curves follow parallel trends and no treatment effect is visually identifiable.

Next, we examine the equivalent short-run effects with parametric DD models. The results are shown in Table B2 (Appendix). In contrast to our main specification, these DD models define the treatment based on child age. Treatment starts at age 9 when children are third graders and we compare the within-child membership development of the treated cohorts against the control cohorts. The models in the first two columns mirror the visual evidence in Figure 4a and focus on Saxony where third graders in the treated school years are compared to third graders in earlier school years before C2SC. By contrast, the next two columns of Table B2 mir-

$$Y_{ics} = \sum_{j=2009}^{2011} \beta_j (Saxony \cdot Cohort_j)_{cs} + \gamma_c + \kappa_s + \varepsilon_{ics}. \quad (4)$$

ror the visual evidence in Figure 4b and focus on third graders in the school years when C2SC was in place (2009/2010, 2010/2011, and 2011/2012). It compares Saxony to the two other states. The uneven columns show the average effect whereas the even columns show event study estimates by child age.[21]

Table B2 again confirms the visual evidence in Figure 4. The point estimates are not only statistically insignificant, but also small and close to zero in size. Furthermore, when estimating event-study regressions in these DD settings (columns [2] and [4] in Table B2), we do not find evidence that the voucher program significantly increased sports club memberships at any of the ages analyzed.[22] The two placebo coefficients provide additional evidence for the common-trend assumptions in these cases.

We can also use the retrospective information on sports club membership at different ages as an outcome in our main specification (equation (2) with municipality fixed effects).[23] Table 3 shows the results for such a specification. It confirms that the C2SC initiative did not significantly increase sports club membership rates among children in the short run. Moreover, when using parents' retrospective responses about their children's sports club membership at different ages, we obtain the same non-significant result (column [3] of Table 3). In conclusion, based on three different, yet related, identification strategies and retrospective information from children as well as their parents, there is little evidence that the C2SC initiative increased sports club membership rates in the short run.

[Insert Table 3 about here]

[21] Column (1) estimates the following model using solely youth who went to primary school in Saxony:

$$Y_{ica} = \beta \cdot Voucher_{ca} + \gamma_c + \mu_a + \varepsilon_{ica},$$

where Y_{ica} denotes the outcome of individual i in cohort c at age a, γ_c are cohort fixed effects, and μ_a stands for a set of age fixed effects. $Voucher_{ca}$ is one if the individual was 9 years or older and went to third grade in school years 2008/09, 2009/10, or 2010/11 (and, therefore, eligible to receive the sports club voucher). $Voucher_{ca}$ leverages the naturally occurring within-child sports club membership probabilities which increase monotonically in a concave manner between ages 5 and 12. Column (3) estimates the following model using only third graders in school years 2008/09, 2009/10, or 2010/11 but in all three states:

$$Y_{isa} = \beta \cdot Voucher_{sa} + \kappa_s + \mu_a + \varepsilon_{isa},$$

where κ_s and μ_a are sets of state and age fixed effects, respectively. $Voucher_{sa}$ is 1 if the individual was 9 years or older and went to third grade in Saxony.

[22] When retrospectively eliciting sports club membership rates across ages, there might be recall bias. For a consistent estimation of the C2SC effect, we assume a common trend in recall biases. More specifically, the assumption for Figure 4a (and columns [1] and [2] in Table B2) is that the recall bias regarding sports club membership before and after age 9 is similar in treated and untreated cohorts in Saxony. The assumption for Figure 4b (and column [3] and [4] in Table B2) is that the recall bias among those who were third graders in 2008/09-2010/11 is similar in Saxony and the other states (regarding pre- and post-age-9 sports club memberships).

[23] This specification then assumes that, at each age, the difference in recall bias between the treatment and control states is similar for treated and untreated cohorts.

There are several potential explanations for the absence of a short-run C2SC effect on membership rates. First, new members would have joined sports clubs irrespective of the C2SC initiative. Figure 4 shows an increase in membership rates by child age, also among non-treated cohorts, providing support for this argument. Second, to affect the consistency of our estimates, recall bias would have to take a non-trivial form. For example, youth may have redeemed the vouchers and joined sports clubs but only for a few weeks, making it more likely for them not to recall their membership several years later. For these individuals, we would also not expect any (long-run) changes in physical activity or health. Third, youth who were already sports club members may have used the voucher to become a member of another sports club. Unfortunately, YOLO only elicits multiple sports club memberships. However, there is some evidence that about one-third of active users redeemed the voucher to experiment with new disciplines (see Figure 5a).

[Insert Figure 5 about here]

5.4 Suggestive Evidence on Mechanisms

Why did the C2SC program fail to significantly increase sports club membership rates? We categorize potential explanations into supply-side and demand-side arguments.

Supply-Side Arguments. Supply-side restrictions would arise if the sports clubs did not have enough capacity for new members. We perform additional analyses to examine the plausibility of this explanation. First, we asked all YOLO respondents who remembered receiving the voucher whether they could redeem it for their desired discipline. Figure 5b shows that 92 percent could redeem the voucher for their desired discipline.

Second, we collected information on the addresses of the 4,381 sports clubs that existed in Saxony in 2008, before the start of C2SC. Based on these addresses, we computed the number of sports clubs in each ZIP code and merged this number with YOLO participants in Saxony. For these individuals, Figure 6a displays the distribution of the number of sports clubs across ZIP codes.[24] Almost 95 percent of YOLO respondents live in ZIP codes with at least six different sports clubs and 11 disciplines; on average, more than 16 sports clubs exist in a ZIP code (median 13). These figures illustrate that, for the large majority of youth, plenty of sports clubs

[24]These respondents live in ZIP codes with about 20,000 inhabitants, on average; 95 percent live in ZIP codes with 9,000 - 40,000 inhabitants (ZIP-code specific population data are taken from https://www.suche-postleitzahl.org/downloads).

existed in the immediate neighborhood.[25] When interpreting these numbers, it is important to note that individuals are not restricted to joining sports clubs in their own ZIP code. Hence, the actual number of sports clubs and available disciplines to choose from is usually larger than the numbers in Figure 6.

[Insert Figure 6 about here]

Third, as another possible barrier, we investigate the self-reported mode of transportation for those youth who indicated that they redeemed the voucher. The descriptive findings, separately by urban and rural ZIP codes are in Figure 5f. As seen, children in rural counties were much more likely to get a ride from their parents (50 vs. 42 percent). However, they were also more likely to walk (30 vs. 23 percent) or bike (34 vs. 30) to the sports club. Finally, not surprisingly, they were significantly less likely to take public transport (9 vs. 22 percent). As these are ex post equilibrium outcomes, it is hard to tell—however—whether a lack of transport was a significant supply side barrier for children. On the one hand, the public transport network is much denser in cities; on the other hand, rural environments make it much safer for children to walk or bike to their sports club.[26]

Fourth, over the course of several years, we personally met with representatives of the Saxony State Sports Association (*Landessportbund Sachsen*) who implemented the policy in cooperation with the Ministry for Education and Cultural Affairs (*Kultusministerium*). In these one-to-one meetings, the representatives informed us that they were not aware of any supply-side constraints at sports clubs in Saxony.

In sum, this suggests that supply-side constraints were unlikely to be a major barrier to take-up.

Demand-Side Arguments. First, it could be that the program targeted the wrong age group. However, as almost half of all third graders in our sample were not sports club members when C2SC started, this is unlikely to be the main explanation (see Figure 4).

[25] Technically, we only observe the number of "divisions" in a sports club, which is a conservative proxy for the number of disciplines: While one discipline is usually organized in the same division, sometimes several disciplines are organized in the same division (e.g., a division entitled "ball games" may include the disciplines of football, handball, and basketball).

[26] Unlike in the United States, "free range parenting" is the social norm in Germany.

Second, stigma could be a reason for incomplete take-up of social benefits (Friedrichsen et al., 2018). As the vouchers were distributed to entire cohorts, stigma is also unlikely to be the main reason for the program's ineffectiveness.

Third, other policies may have confounded the demand side. We have carefully checked the legislation in the relevant federal states and found no such policies in Saxony or the neighboring states.[27] The official school curricula show that physical education hours did not change in any of the three states at the time. In Saxony, students in grades one to ten have three hours of physical education per week. The requirements in Brandenburg and Thuringia are almost identical.[28]

Fourth, a lack of parental information may have created access barriers. These might be particularly relevant for children from disadvantaged backgrounds. However, several C2SC features could have mitigated such issues: Parents received information packages which described C2SC, explained how to redeem the vouchers, and also listed all local sports clubs along with the disciplines they offered (see Section 2.2). Further, because entire cohorts were treated, it is very likely that parents heard about the program from other parents, and also from their children. Nevertheless, a lack of information or support from parents cannot be ruled out as a take-up barrier.

Fifth, besides a lack of parental information, a lack of involvement and encouragement from parents might be driving take-up barriers. Table B3 in the Appendix compares the characteristics of parents of children who are vs. those who are not sports club members. This is simply a descriptive exercise and considers all children in all cohorts and all states. As seen, it is very clear that children of better educated parents and parents who are sports club members themselves are significantly more likely to become sports club members themselves. Further, we surveyed parents' attitudes about exercising in general in the parent questionnaire (see notes to Table B3 for details). Responses are on a scale from 1 (totally disagree) to 7 (totally agree). Table B3 shows that differences in parental attitudes are highly significant for all 11 questions. Parents whose children are sports club members are more likely to agree that, for them, exercising is something that is integrated into their weekly routine, that is a habit, that they do without a lot of effort and willpower, and that they have been doing for a long time. Next, we conduct

[27] We are also not aware of policies affecting the supply-side, i.e., sports clubs. As discussed in footnote 5, the federal "Educational Package", which covered membership fees for welfare recipients, came into effect on April 1, 2011, after the third C2SC cohorts' deadline to redeem the voucher. Further, the Educational Package affected welfare recipients in all states.

[28] In Thuringia, students in grades 1 and 2 only have two hours of physical education.

a similar exercise but focus on the treated cohorts in Saxony and differentiate between children who redeemed vs. those who did not redeem the voucher. This time we use responses from the youth questionnaire. The findings in Table B4 of the Appendix clearly show that children who redeemed the voucher are significantly more likely to have art and newspapers at home, are on the highest educational track, and were already sports club members between ages 4 and 7 (before the treatment). These children were also significantly more likely to have taken music lessons between ages 4 and 7. All these factors remain statistically significant in a multivariate regression framework. Interestingly, and reinforcing our supply-side discussion, the number of sports clubs in the ZIP code does *not* statistically differ between the two groups of children.

Sixth, although membership fees were waived, additional monetary and non-monetary costs are typically associated with sports club memberships and may prevent take-up. In particular, C2SC did not cover costs for equipment. Further, the vouchers only covered the membership fees for a single year. However, Figures 5c and d show that monthly membership fees are not the main barrier to take-up for most families. Only three percent of youth who redeemed the voucher responded that membership fees would have been unaffordable without the voucher. Consistently, a quarter of active users said that their parents were glad to be able to save money thanks to the vouchers. However, for parents, sports club memberships sometimes carry large non-monetary costs; for instance, when parents have to take their children to the sports facilities. Further, on weekends, parents are often asked to volunteer in tournaments or in matches against other teams.

In sum, (future) financial costs, indirect costs, and a lack of parental involvement might explain why few children in the target group redeemed the vouchers. We cannot and do not attempt to distinguish these diverse potential demand-side explanations but interpret the evidence as suggestive that supply-side constraints were not the major driving force. Conversely, we find clear evidence that children who redeemed the vouchers were significantly more likely to come from households with art and newspapers at home and to be on a higher educational track. It is also clear that they were already sports club members prior to the treatment and received music lessons. Moreover, their parents are significantly more likely to hold positive attitudes about exercising and to see it as an integral part of their daily routine. Hence parental involvement, attitudes and encouragement appear to be very important driving forces for children's voucher take-up.

6 Sensitivity Analyses

6.1 Alternative Specifications, Placebos, and Further Outcomes

The first set of robustness checks addresses alternative model specifications, in particular, alternative sample restrictions, alternative assignments of the treatment group, and alternative ways of dealing with control variables. Table 4 shows results for our preferred specification and all six outcomes, where we vary the cohorts included in the sample (columns [1] to [4]), omit either Brandenburg or Thuringia from the sample (columns [5] and [6]), include individuals who lived in other states when they were third graders (column [7]), and only consider individuals with matching registry information (regarding gender, nationality, and age) and who spent at least ten minutes on the survey (column [8]). In column (9) of Table 4, to consider potential spillover effects between siblings, we exclude respondents from the pre-treatment cohorts in all three states who reported that their siblings also received an invitation to participate in YOLO. In column (10), we exclude youth with an older sibling.

To test for whether possibly misreported migration between states and the first and third grade matters, when we define the treatment group, we use information on where the child attended first grade (column [11]) and use the current state of residence (column [12]).[29] Column (13) controls for socio-demographics[30], while column (14) weights the sample based on the probability of an individual participating in the YOLO survey, which we derived from the registry information.[31] The next set of robustness checks focuses on placebo tests. First, we omit Saxony from the sample and assume that Brandenburg was the treatment state and Thuringia the control state (see column [14] of Table 4). All point estimates are small in size, also the first three on awareness and utilization, and not statistically different from zero.

[Insert Table 4 about here]

Second, Figure 7a plots point estimates from our preferred model along with 95 percent confidence bands. However, here we use outcomes that either cannot plausibly be affected

[29] Note that geographic mobility in Germany is much lower than in the US. In YOLO, 92 percent of youth also spent the third grade in the state of their current residency, about a decade later. According to the representative SOEP, in a given year, less than one percent of all respondents move (Jürges et al., 2011; Goebel et al., 2015).

[30] These control variables include the binary variables female, has siblings, born in Germany, parent not born in Germany, newspaper at home, art at home, academic track, sports club aged 4-7, and music lessons aged 4-7 (see Table A2, Appendix).

[31] More specifically, we apply inverse probability weights, where the probability to participate in the YOLO survey is predicted by a Probit model. This Probit model uses registry information and includes a fully interacted set of state, gender, year of birth, and German nationality dummies.

by the treatment, such as gender, the number of siblings, or whether the child was born in Germany—or outcomes that are very unlikely to be affected by vouchers, such as *newspaper at home*, *art at home*, or *academic track*. We also use sports club participation during childhood ages prior to the treatment (between age four and seven) as a placebo outcome. As seen, the effect on none of these outcomes is statistically significant, but all point estimates are very close to zero. All these robustness checks confirm that the C2SC initiative increased awareness and take-up, but had no long-run effect on physical activity and being overweight.

Figure 7b displays treatment effects for alternative outcomes, such as different measures of physical activity or attitudes regarding sports. We also measure other health behaviors such as drinking or smoking. In conclusion, C2SC affected none of the sport, health, overweightness, smoking, or alcohol-consumption outcomes. This is in line with our main findings and supports our conclusion that the C2SC initiative did not improve health or health-behaviors in the long-run.

[Insert Figure 7 about here]

In addition to weighting the regressions with the respondents' probability of participating in YOLO (column [13] of Table 4), Table A6 (Appendix) shows that individuals in treated cohorts in Saxony are just as likely as everybody else to participate in YOLO. Using only registry data, we re-run our main DD model in equation (2), but use the dummy "YOLO participation" as an outcome and assign the treatment based on the date of birth. As seen in Table A6, the small and insignificant point estimate provides evidence that survey participation is not significantly related to treatment status. Table B5 (Appendix) employs 12 different methods of statistical inference (including different levels of clustering and different wild cluster bootstrap procedures). None of the estimates lead us to revise our main conclusion; all 36 estimated coefficients for the first three outcomes remain highly significant, and all 36 coefficients for the other three outcomes remain insignificant.

6.2 Parents' Information

We also invited parents to fill out a similar survey, which we electronically linked to their children's responses (see Section 3). Now, we leverage this information to check whether the results are robust to using parents' responses. Table 5 shows the results. The first two columns use the children's responses as a benchmark and the third column uses parents' responses. We always

show our preferred DD estimates.[32] As seen, using parents' responses regarding whether their child received and redeemed the voucher about a decade ago, the estimates remain highly significant and are very similar to the estimates obtained using children's responses. If anything, the program is slightly better known among children (comparing columns [2] and [3] in Table 5), while the utilization effect is slightly larger when using the parents' responses.

[Insert Table 5 about here]

6.3 Synthetic Control Group

Although the pre-treatment trends in Figure 1 are fairly similar, as another test, we follow Abadie et al. (2010) and construct a synthetic control group with the same pre-treatment outcome trends. More specifically, we re-weight the municipalities in the control states such that the two pre-treatment cohorts have the same mean as the treatment state of Saxony with respect to *member of sports club*, *weekly hours of sport*, and *overweight*.[33] Based on this synthetic control group approach, neither the regression results in Table 6 nor the visual inspections in Figure B2 (Appendix) provide evidence that the C2SC initiative increased membership rates and the amount of physical exercise, or reduced overweight rates.

[Insert Table 6 about here]

This finding is further reinforced when we plot the differences between treatment and control states both for the main sample (Panel A) and the synthetic control group approach (Panel B) in the Appendix in Figure B3. As seen all point estimates are very close to zero and almost no trending is discernible. Further, all confidence intervals largely overlap with zero and the point estimate signs alternate between being positive and negative.

6.4 Statistical Power

As a final exercise, we conduct two types of power analyses. The results are in the Appendix in Tables B6 and B7. First, for our main outcomes and four different models as indicated in the first column, Table B6 shows the one-sided limits of 90 percent (column 2) and 95 percent (column 3) confidence intervals. For example, the model in the first row of Panel A in column

[32] Note that we did not ask parents about their children's current sports activities or weight.
[33] The notes for Table 6 provide further details on the construction of the synthetic control group.

(3) indicates that increases in sports club membership rates by more than 2.8 percentage points as a result of C2SC lie outside the 90 percent confidence interval. The models in Panel B and C can exclude increases of 2.7 and 3.3 percentage points, respectively. Note that, as a precaution, we excluded the last surveyed cohort (2011/2012) from our main models. This cohort was not treated and is thus a control group, but it could be hypothesized that disappointment effects might confound the main effects (see footnote 9). However, to increase our statistical power, we re-include that cohort in the models in Panel D. As seen, here we can exclude increases in membership rates by more than 0.5 percentage points. Similarly, this model also has the tightest bounds to exclude increases in weekly hours of sport of more than 0.005 as well as decreases in overweight of more than 0.008 percentage points. In the next subsection, we will use administrative health examination data which can exclude overweight decreases by more than 0.07 percentage points and motor disorder decreases by more than 0.43 percentage points at the 90 percent statistical certainty level.

Second, Table B7 shows the results of 36 power simulations with 1,000 replications each (see notes to the table for specific details). Again, the first column indicates our model specifications and outcome variables. Columns (1) to (3) simulate the ability of our models to identify one-sided treatment effects for 3 to 5 units with 90 percent statistical certainty. Specifically, we artificially add pseudo treatment effects to our models—for example, increases of 0.3, 0.4. as well as 0.5 hours of sport when our outcome of interest is *weekly hours of sports*—and then run the models 1000 times per specification. The numbers in Table B7 indicate how many of these 1,000 replications correctly reject the null hypothesis of no treatment effects at the 10 percent significance level in one-sided tests. For example, the simulation in the third row of Panel A in column (2) indicates that our baseline model in Table 2 can identify statistically significant decreases in the overweight rate of 4 percentage points in 88 percent of all cases. Rules of thumb of power analyses suggest that 80 percent is a solid threshold for well-powered models (cf. Griffith and Feyman, 2020). As seen, all our models exceed this threshold in column (3). The model in Panel D with extended power through the inclusion of the 2011/2012 cohort also reaches this threshold for all three outcomes and an artificial treatment effect of 4 units (column 2). It is also noteworthy that neither the method of synthetic controls (Panel C) nor the inclusion of additional controls (Panel B) appear to increase our statistical power substantially.

7 Additional Evidence Based on Official Health Examination Data

In this section, we use supplemental school health examination data, obtained from one of the 13 counties in Saxony. Paragraph 26a of the Saxony School Law (*Sächsisches Schulgesetz–SächsSchulG*) stipulates that the Public Health Service (*Öffentlicher Gesundheitsdienst*), together with the principal, the teachers, parents, and school children, carry out or participate in School Health Care (*Schulgesundheitspflege*). An integral part of this is the physical examination carried out by a state-employed physician for all children in the second and sixth grade of all schools to identify and prevent diseases and development disorders among children (*Schulreihenuntersuchung*). The examinations are mandatory and all children are legally obligated to attend school in Germany; home schooling does not exist. The state-employed physician, with the help of physician assistants, measure the height and weight of all children, check immunizations, and test for development or motor skill disorders. They write recommendations for parents, schedule follow-up visits, and refer children to specialists or therapists if necessary.

We obtained school medical examination data for *the universe of* sixth graders in one county. The county has a population of about 300,000 inhabitants, where ten percent are children and adolescents aged 6-18 (Statistisches Landesamt, Freistaat Sachsen, 2020). The data include health examination data of sixth graders in school years 2009/10, 2010/11 vs. 2012/13, 2013/14. Hence, the first two cohorts were not affected by C2SC, whereas the last two cohorts were part of the second and third round of the C2SC initiative, when the children were third graders in school years 2009/10 and 2010/11. That is, the treated cohorts received the vouchers at the end of January 2010 and January 2011 and were physically examined between mid-March and mid-July 2012 and 2013, slightly more than two years after the beginning of the C2SC initiative.[34]

Because all data represent objective diagnoses by a small number of state-employed physicians, measurement errors should play a negligible role (Salm and Schunk, 2012). In terms of outcome variables, we rely on the coding used by the state-employed physician. For example, the standard body mass index (BMI) cut-offs to determine overweightness and obesity for adults do not apply to children. Guidelines determine overweightness and obesity in per-

[34] We only consider students who were in the sixth grade at the time of the examination and who started primary school in 2004, 2005, 2007, and 2008 (that is, we exclude students who skipped or repeated a grade). We also focus on children in "regular" schools and exclude those in special schools for the disabled or those with specific disorders. This is a census of all children. When children are sick on the day of the examination, there is a follow-up and they will be examined on the next possible date. Moreover, there are basically no missings in the data; we have missing for BMI per year for a cohort of about 2.5 thousand children.

centiles relative to the age in months of the child. The state-employed physicians in Saxony follow the guidelines of the "German Working Group on Child and Youth Obesity" (Arbeitsgemeinschaft Adipositas im Kindes- und Jugendalter (AGA) , 2019), who refer to Kromeyer-Hauschild et al. (2001, 2015). We use the diagnoses recorded by the state-employed physicians. The guidelines define children as overweight if they fall between the 90^{th} and 97^{th} percentile of the age-specific BMI distribution, as obese if they exceed the 97^{th} percentile, and as underweight if they fall below the 10^{th} percentile of the age-specific BMI distribution. The major advantage of these data is the objective nature of the health and motor skills measures.

Table C1 (Appendix) shows that the data include 6,794 observations. Half of all children are female and the average age is 12.1 years (145 age-months as recorded in the data). The average height is 154 centimeters (5 foot) and the average weight is 46 kilograms (101 pounds). The calculated BMI varies between 11 and 42. A total of 11 percent of all children are overweight and underweight, respectively, and six percent are obese. Eight percent had hypertension, 11 percent a poor posture, 0.4 percent motor skill disorders, and 3.6 percent emotional disorders.

Next, we compare the objective health measures of the two treated cohorts against the two control cohorts. For causal inference of the C2SC initiative, we have to assume the absence of cohort effects. Consequently, we interpret the estimates below with caution.

[Insert Figure 8 about here]

Figures 8a to d compare the share of sixth graders who are (a) obese or (b) overweight, and who have (c) motor skill disorders and (d) emotional disorders in the treatment and control cohorts. Because our main findings do not deliver any evidence of significant long-term effects on physical activity, we hypothesize that there will not be much evidence of health improvements either. As seen, Figure 8 confirms these priors. We obtain the same finding when regressing the objective health outcomes on an indicator for the treatment cohorts and a set of pre-determined control variables in Table C2 (Appendix). Overall, the supplemental school health examination data confirm our findings based on the YOLO survey.

8 Discussion and Conclusion

Integrating the habit of regular physical activity into our daily lives is generally considered a worthwhile objective (Centers for Disease Control and Prevention, 2011; European Commis-

sion, 2016; World Health Organization, 2018). Empirical studies from different disciplines have linked physical activity to a wide array of positive outcomes, including lower risks of chronic diseases, better sleep, and normal development of children (Piercy et al., 2018). Economic studies provide evidence that physical activity is associated with lower obesity rates (Cawley et al., 2013), improved educational performance (Stevenson, 2010; Fricke et al., 2018)[35], and better labor market outcomes (Lechner, 2009; Stevenson, 2010). It is not only individuals who may benefit from regular exercise, but also society as a whole. People who exercise regularly might be more resilient to stress (Childs and de Wit, 2014), show better life skill development (Gould and Carson, 2008), higher levels of trust (Schüttoff et al., 2018), and have a higher degree of social capital (Di Bartolomeo and Papa, 2019). At the same time, a comprehensive literature suggests that habit formation plays a key role in adopting a healthier lifestyle; the earlier in life people start being physically active, the stronger the long-term effects (Hallal et al., 2006; Lally and Gardner, 2013; Belot et al., 2016; Loewenstein et al., 2016).

This paper evaluates a policy aimed at increasing the physical activity levels of third graders in the long term. The hope and the policy's intention was to encourage children, especially those from disadvantaged backgrounds, to adopt the habit of exercising regularly in order to improve physical fitness. However, in order not to discriminate against anyone, vouchers were distributed among *all* 33,000 third graders in the German state of Saxony in January 2009. The treatment was repeated twice with the next two cohorts in January 2010 and 2011. Redeemed vouchers provided free sports club membership of up to one year. The findings produced by this policy initiative differ from existing evidence as the initiative represents one of the very few quasi-experimental settings explicitly targeting children. What is more, most existing field experiments focus on gyms and "only" strive to change adults' habits for a few weeks or months, while this treatment lasted for an entire year. Moreover, the empirical setting allows us to estimate long-run effects seven to nine years post-intervention.

Drawing on a unique register-based survey and difference-in-differences approaches, our findings demonstrate that those who were treated about a decade ago still recall the initiative today. They also redeemed the vouchers at significantly higher rates. However, the program was not effective in raising levels of regular physical activity and in reducing overweight among youth. We find neither significant nor suggestive evidence that the policy attracted a substantial share of new members. We discuss several potential explanations for the ineffec-

[35]However, there is also some evidence of negative effects (Golsteyn et al., 2020).

tiveness of the program. Supply-side constraints are unlikely to be a driving force as most children could choose from many sports clubs in close proximity. Rather, demand-side explanations such as additional monetary costs (e.g., equipment) and high non-monetary costs (e.g., parental time) are the most plausible explanations for why the program did not increase sports club memberships. We find that it was primarily children from advantaged socio-demographic backgrouds who had already been sports club members who redeemed the vouchers, questioning the effectiveness of untargeted voucher programs (see also Schwerdt et al., 2012). In line with our finding that the voucher program was ineffective in encouraging a significant share of children to become members of sports clubs and exercise regularly, we find no change in health behaviors or objective health using official health examination data either in the short- or long-run. Empirically evaluating which measures are effective in encouraging youth to integrate physical activity into their daily lives will certainly remain a fruitful and highly relevant research field across the social science disciplines.

References

Abadie, A., A. Diamond, and J. Hainmueller (2010). Synthetic control methods for comparative case studies: Estimating the effect of California's tobacco control program. *Journal of the American Statistical Association 105*(490), 493–505.

Abadie, A., G. W. Imbens, S. Athey, and J. M. Wooldridge (2017). When should you adjust standard errors for clustering? *NBER Working Paper 24003*.

Akee, R., E. Simeonova, W. Copeland, A. Angold, and E. J. Costello (2013). Young adult obesity and household income: Effects of unconditional cash transfers. *American Economic Journal: Applied Economics 5*(2), 1–28.

Anderson, M. L. and D. A. Matsa (2011). Are restaurants really supersizing America? *American Economic Journal: Applied Economics 3*(1), 152–188.

Angelucci, M., S. Prina, H. Royer, and A. Samek (2020). When incentives backfire: Spillover effects in food choice. *American Economic Journal: Economic Policy 11*(4), 66–95.

Arbeitsgemeinschaft Adipositas im Kindes- und Jugendalter (AGA) (2019). *Definition der Adipositas*. https://aga.adipositas-gesellschaft.de, last accessed on May 4, 2019.

Augurzky, B., T. K. Bauer, A. R. Reichert, C. M. Schmidt, and H. Tauchmann (2018). Habit formation, obesity, and cash rewards. Ruhr Economic Papers 750.

Babcock, P., K. Bedard, G. Charness, J. Hartman, and H. Royer (2015). Letting down the team? Social effects of team incentives. *Journal of the European Economic Association 13*(5), 841–870.

Belot, M., J. James, and P. Nolen (2016). Incentives and children's dietary choice: A field experiment in primary schools. *Journal of Health Economics 50*, 213–229.

Breuer, C. and S. Feiler (2015). Sportvereine in Deutschland – ein Überblick. In C. Breuer (Ed.), *Sportentwicklungsbericht 2013/2014. Analyse zur Situation der Sportvereine in Deutschland*, pp. 1–42. Köln: Sportverlag Strauß.

Breuer, C., S. Feiler, and P. Wicker (2015). Sport clubs in Germany. In C. Breuer, R. Hoekman, S. Nagel, and H. van der Werff (Eds.), *Sport clubs in Europe*, pp. 187–208. Switzerland: International Publishing Springer.

Breuer, C., R. Hoekman, S. Nagel, and H. v. d. Werff (Eds.) (2015). *Sport Clubs in Europe: A Cross-National Comparative Perspective*. Sports Economics, Management and Policy. Springer International Publishing.

Carrera, M., H. Royer, M. Stehr, and J. Sydnor (2018). Can financial incentives help people trying to establish new habits? Experimental evidence with new gym members. *Journal of Health Economics 58*(C), 202–214.

Carrera, M., H. Royer, M. F. Stehr, and J. R. Sydnor (2020). The structure of health incentives: Evidence from a field experiment. *Management Science 66*(5), 1783–2290.

Cawley, J., D. Frisvold, and C. Meyerhoefer (2013). The impact of physical education on obesity among elementary school children. *Journal of Health Economics 32*(4), 743 – 755.

Cawley, J., C. Meyerhoefer, and D. Newhouse (2007). The impact of state physical education requirements on youth physical activity and overweight. *Health Economics 16*(12), 1287–1301.

Centers for Disease Control and Prevention (2011). *Strategies to prevent obesity and other chronic diseases: The CDC Guide to strategies to increase physical activity in the community*. Atlanta, GA: U.S. Department of Health and Human Services.

Childs, E. and H. de Wit (2014). Regular exercise is associated with emotional resilience to acute stress in healthy adults. *Frontiers in Physiology 5*, 161.

Courtemanche, C. (2009). Rising cigarette prices and rising obesity: Coincidence or unintended consequence? *Journal of Health Economics 28*(4), 781–798.

Courtemanche, C. (2011). A silver lining? The connection between gasoline prices and obesity. *Economic Inquiry 49*(3), 935–957.

Courtemanche, C., G. Heutel, and P. McAlvanah (2015). Impatience, incentives and obesity. *The Economic Journal 125*(582), 1–31.

Currie, J., S. DellaVigna, E. Moretti, and V. Pathania (2010). The effect of fast food restaurants on obesity and weight gain. *American Economic Journal: Economic Policy 2*(3), 32–63.

Currie, J. and F. Gahvari (2008). Transfers in cash and in-kind: Theory meets the data. *Journal of Economic Literature 46*(2), 333–383.

Cutler, D., E. L. Glaeser, and J. M. Shapiro (2003). Why have Americans become more obese? *Journal of Economic Perspectives 17*(3), 93–118.

DellaVigna, S. and U. Malmendier (2006). Paying not to go to the gym. *American Economic Review 96*(3), 694–719.

Di Bartolomeo, G. and S. Papa (2019). The effects of physical activity on social interactions: The case of trust and trustworthiness. *Journal of Sports Economics 20*(1), 50–71.

Dragone, D. and N. R. Ziebarth (2017). Non-separable time preferences, novelty consumption and body weight: Theory and evidence from the East German transition to capitalism. *Journal of Health Economics 51*, 41 – 65.

Dubois, P., R. Griffith, and A. Nevo (2014). Do prices and attributes explain international differences in food purchases? *American Economic Review 104*(3), 832–867.

Dunn, R. A. (2010). The effect of fast-food availability on obesity: An analysis by gender, race, and residential location. *American Journal of Agricultural Economics 92*(4), 1149–1164.

European Commission (2016). *Monitoring the activities of the EU Platform for Action on Diet, Physical Activity and Health. Annual Report 2016.* Brussels: European Commission.

Felfe, C. and R. Lalive (2018). Does early child care affect children's development? *Journal of Public Economics 159(C)*, 33 – 53.

Fletcher, J. M., D. E. Frisvold, and N. Tefft (2010). The effects of soft drink taxes on child and adolescent consumption and weight outcomes. *Journal of Public Economics 94*(11-12), 967 – 974.

Fricke, H., M. Lechner, and A. Steinmayr (2018). The effects of incentives to exercise on student performance in college. *Economics of Education Review 66(C)*, 14–39.

Frieden, T. R., W. Dietz, and J. Collins (2010). Reducing childhood obesity through policy change: Acting now to prevent obesity. *Health Affairs 29*(3), 357–363.

Friedrichsen, J., T. König, and R. Schmacker (2018). Social image concerns and welfare take-up. *Journal of Public Economics 168*, 174–192.

German Olympic Sports Confederation (2011). Sportgutscheine für Sachsens Drittklässler. News, DOSB-Presse.

German Olympic Sports Confederation (2017). *Mitgliederentwicklung in Sportvereinen 2000 bis 2015. Bestand, Veränderungen und Perspektiven*. Frankfurt am Main: German Olympic Sports Confederation.

Goebel, J., M. M. Grabka, S. Liebig, M. Kroh, D. Richter, C. Schröder, and J. Schupp (2019). The German Socio-Economic Panel (SOEP). *Jahrbücher für Nationalökonomie und Statistik 239*(2), 345–360.

Goebel, J., C. Krekel, T. Tiefenbach, and N. Ziebarth (2015). How natural disasters can affect environmental concerns, risk aversion, and even politics: Evidence from Fukushima and three European countries. *Journal of Population Economics 28*(4), 1137–1180.

Golsteyn, B. H. H., M. W. J. Jansen, D. H. H. Van Kann, and A. M. C. Verhagen (2020). Does stimulating physical activity affect school performance? *Journal of Policy Analysis and Management 39*(1), 64–95.

Gould, D. and S. Carson (2008). Life skills development through sport: current status and future directions. *International Review of Sport and Exercise Psychology 1*(1), 58–78.

Graf, S. and M. Cecchini (2019). Current and past trends in physical activity in four OECD countries: Empirical results from time use surveys in Canada, France, Germany and the United States. *OECD Health Working Papers 112*.

Griffith, K. N. and Y. Feyman (2020). Amplifying the noise: The dangers of post hoc power analyses. *Journal of Surgical Research forthcoming*.

Griffith, R., R. Lluberas, and M. Lührmann (2016). Gluttony and sloth? Calories, labor market activity and the rise of obesity. *Journal of the European Economic Association 14*(6), 1253–1286.

Grossman, M., E. Tekin, and R. Wada (2014). Food prices and body fatness among youths. *Economics & Human Biology 12*, 4–19.

Hainmueller, J. (2012). Entropy balancing for causal effects: A multivariate reweighting method to produce balanced samples in observational studies. *Political Analysis 20*(1), 25–46.

Hallal, P. C., C. G. Victora, M. R. Azevedo, and J. C. K. Wells (2006). Adolescent physical activity and Health: A Systematic Review. *Sports Medicine 36*(12), 1019–1030.

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

Jeitschko, T. D. and R. A. Pecchenino (2006). Do you want fries with that? An exploration of serving size, social welfare, and our waistlines. *Economic Inquiry 44*(3), 442–450.

Jürges, H., S. Reinhold, and M. Salm (2011). Does schooling affect health behavior? Evidence from the educational expansion in Western Germany. *Economics of Education Review 30*(5), 862 – 872. Special Issue on Education and Health.

Just, D. R. and J. Price (2013). Using incentives to encourage healthy eating in children. *Journal of Human Resources 48*(4), 855–872.

Kesternich, I., B. Siflinger, J. P. Smith, and J. K. Winter (2015). Individual behavior as a pathway between early-life shocks and adult health: Evidence from hunger episodes in post-war Germany. *The Economic Journal 125*(588), F372–F393.

Kreissportbund Landkreis Leipzig (2019). *Projekt: 'KOMM! Engagier Dich'*. https://www.ksb-ll.de/kinder-jugendsport-leipzig/komm-in-den-sportverein-landkreis-leipzig/, last accessed on May 7, 2019.

Kromeyer-Hauschild, K., M. A., and M. Wabitsch (2015). Referenzwerte für den Body-Mass-Index für Kinder, Jugendliche und Erwachsene in Deutschland. Anpassung der AGA-BMI-Referenz im Altersbereich von 15 bis 18 Jahren. *Adipositas 149*, 807–818.

Kromeyer-Hauschild, K., M. Wabitsch, D. Kunze, and et al. (2001). Perzentile für den Body-Mass-Index für das Kindes- und Jugendalter unter Heranziehung verschiedener deutscher Stichproben. *Monatsschrift Kinderheilkunde 9*, 123–127.

Lakdawalla, D. and T. Philipson (2009). The growth of obesity and technological change. *Economics and Human Biology 7*(3), 283–293.

Lally, P. and B. Gardner (2013). Promoting habit formation. *Health Psychology Review 7*(sup1), S137–S158.

Landessportbund Sachsen (2019). *Kampagne 'KOMM! Engagier Dich' Engagementförderung im sächsischen Sport*. https://www.sport-fuer-sachsen.de/sport-und-gesellschaft/engagementfoerderung/, last accessed on May 4, 2019.

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

Loewenstein, G., J. Price, and K. Volpp (2016). Habit formation in children: Evidence from incentives for healthy eating. *Journal of Health Economics 45*, 47–54.

Peter, F. H. and C. K. Spiess (2016). Family Instability and Locus of Control in Adolescence. *The B.E. Journal of Economic Analysis & Policy 16*(3), 1439–1471.

Piercy, K. L., R. P. Troiano, R. M. Ballard, S. A. Carlson, J. E. Fulton, D. A. Galuska, S. M. George, and R. D. Olson (2018). The physical activity guidelines for Americans. *JAMA 320*(19), 2020–2028.

Prentice-Dunn, H. and S. Prentice-Dunn (2012). Physical activity, sedentary behavior, and childhood obesity: A review of cross-sectional studies. *Psychology, Health & Medicine 17*(3), 255–273.

Prina, S. and H. Royer (2014). The importance of parental knowledge: Evidence from weight report cards in Mexico. *Journal of Health Economics 37*, 232 – 247.

Reichert, A. R. (2015). Obesity, weight loss, and employment prospects: Evidence from a randomized trial. *Journal of Human Resources 50*(3), 759–810.

Roodman, D., M. Ø. Nielsen, J. G. MacKinnon, and M. D. Webb (2010). Fast and wild: Bootstrap inference in stata using boottest. *The Stata Journal 19*(1), 4–60.

Royer, H., M. Stehr, and J. Sydnor (2015). Incentives, commitments, and habit formation in exercise: Evidence from a field experiment with workers at a fortune-500 company. *American Economic Journal: Applied Economics 7*(3), 51–84.

Salm, M. and D. Schunk (2012, 12). The Relationship between Child Health, Developmental Gaps, and Parental Education: Evidence from Administrative Data. *Journal of the European Economic Association 10*(6), 1425–1449.

Schüttoff, U., T. Pawlowski, P. Downward, and M. Lechner (2018). Sports participation and social capital formation during adolescence. *Social Science Quarterly 99*(2), 683–698.

Schwerdt, G., D. Messer, L. Woessmann, and S. C. Wolter (2012). The impact of an adult education voucher program: Evidence from a randomized field experiment. *Journal of Public Economics 96*(7-8), 569 – 583.

Siedler, T., J. Schupp, C. K. Spiess, and G. G. Wagner (2009). The German Socio-Economic Panel (SOEP) as reference data set. *Schmollers Jahrbuch 129*(2), 367–374.

Statistisches Bundesamt (2020). *GENESIS-ONLINE: Die Datenbank des Statistischen Bundesamtes*. https://www-genesis.destatis.de/genesis/online, last accessed on June 26, 2020.

Statistisches Landesamt, Freistaat Sachsen (2020). *Bevölkerungsstand des Freistaatest Sachsen nach Alter und Geschlecht*. https://www.statistik.sachsen.de/download/100_Berichte-A/A_I_3_j18_SN.pdf, last accessed on March 26, 2020.

Stevenson, B. (2010). Beyond the classroom: Using Title IX to measure the return to high school sports. *The Review of Economics and Statistics 92*(2), 284–301.

Strulik, H. (2014). A mass phenomenon: The social evolution of obesity. *Journal of Health Economics 33(C)*, 113–125.

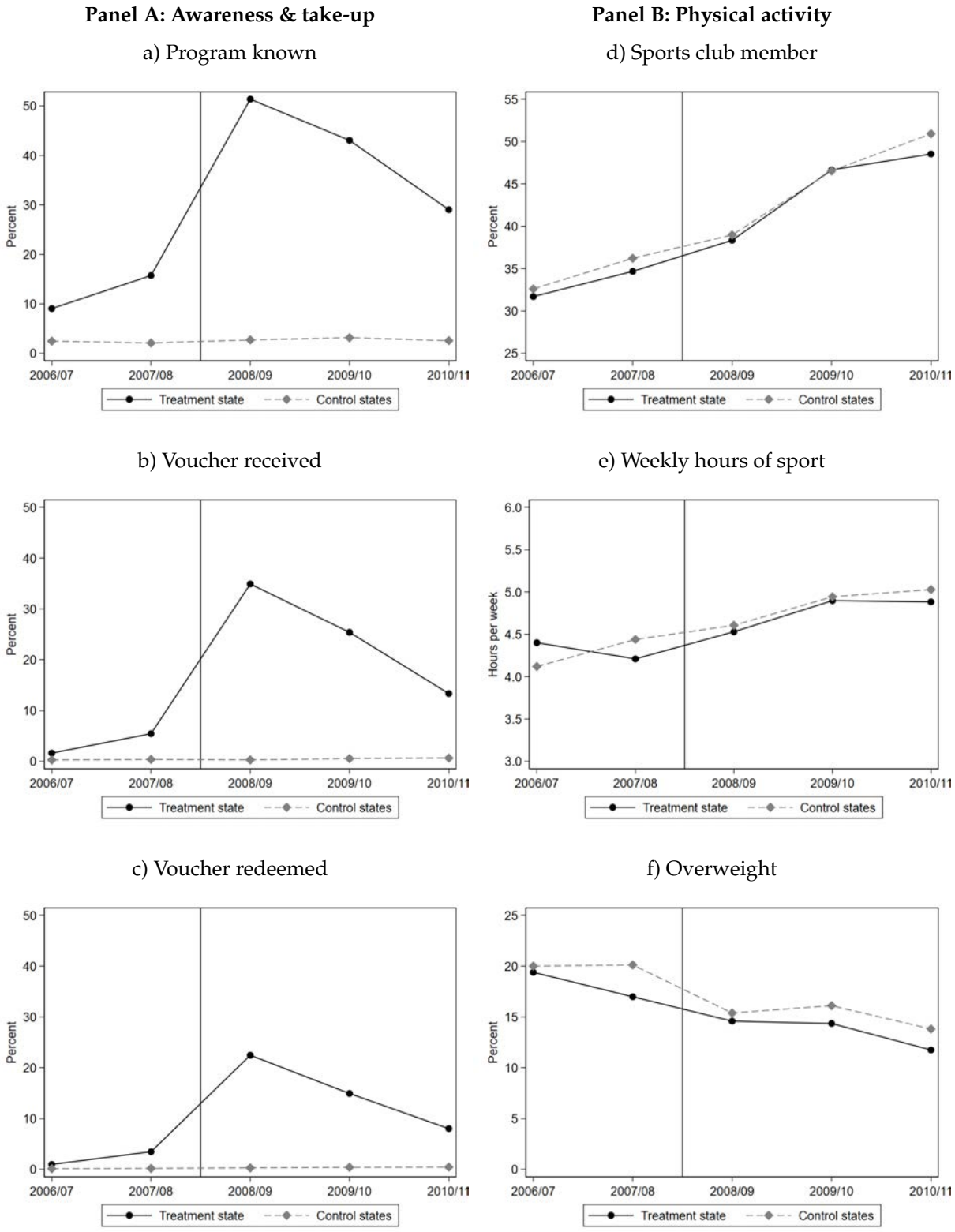
World Health Organization (2017). *Report of the Commission on Ending Childhood Obesity. Implementation plan: Executive summary*. Geneva: World Health Organization. Licence: CC BY-NC-SA 3.0 IGO, https://www.who.int/end-childhood-obesity/publications/echo-plan-executive-summary/en/, last accessed on March 26, 2020.

World Health Organization (2018). *Physical Activity*. https://www.who.int/news-room/fact-sheets/detail/physical-activity last accessed March 26, 2020.

World Health Organization (2020). *Global Strategy on Diet, Physical Activity and Health: Childhood Overweight and Obesity*. https://www.who.int/dietphysicalactivity/childhood/en/, last accessed on March 26, 2020.

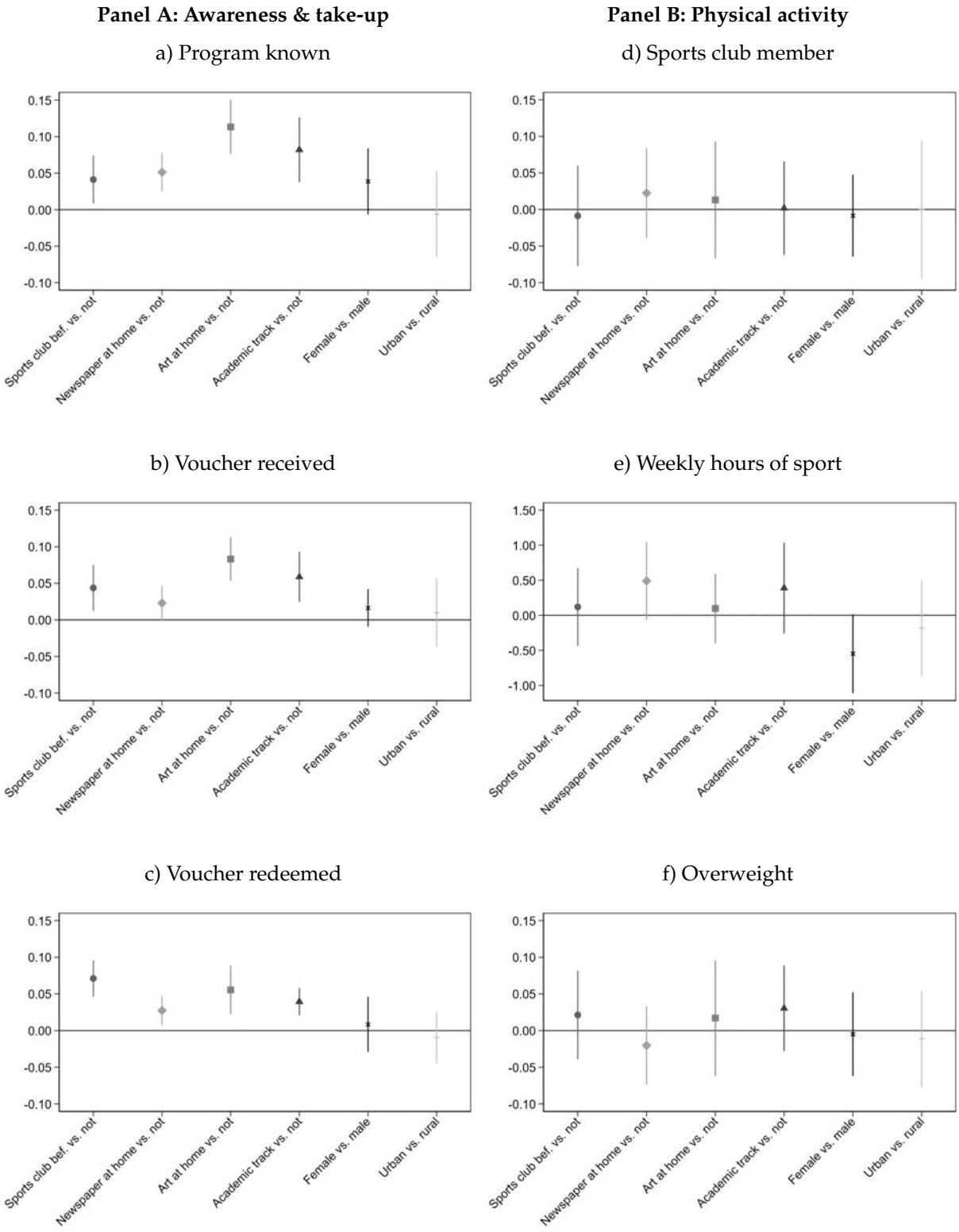
Züchner, I. and B. Arnoldt (2012). Sport von Kinder und Jugendlichen in Ganztagsschule und Sportverein. *Körper, Bewegung und Schule 3*(6), 333–383.

Figure 1: Development of Outcome Variables in Treatment and Control States across Cohorts



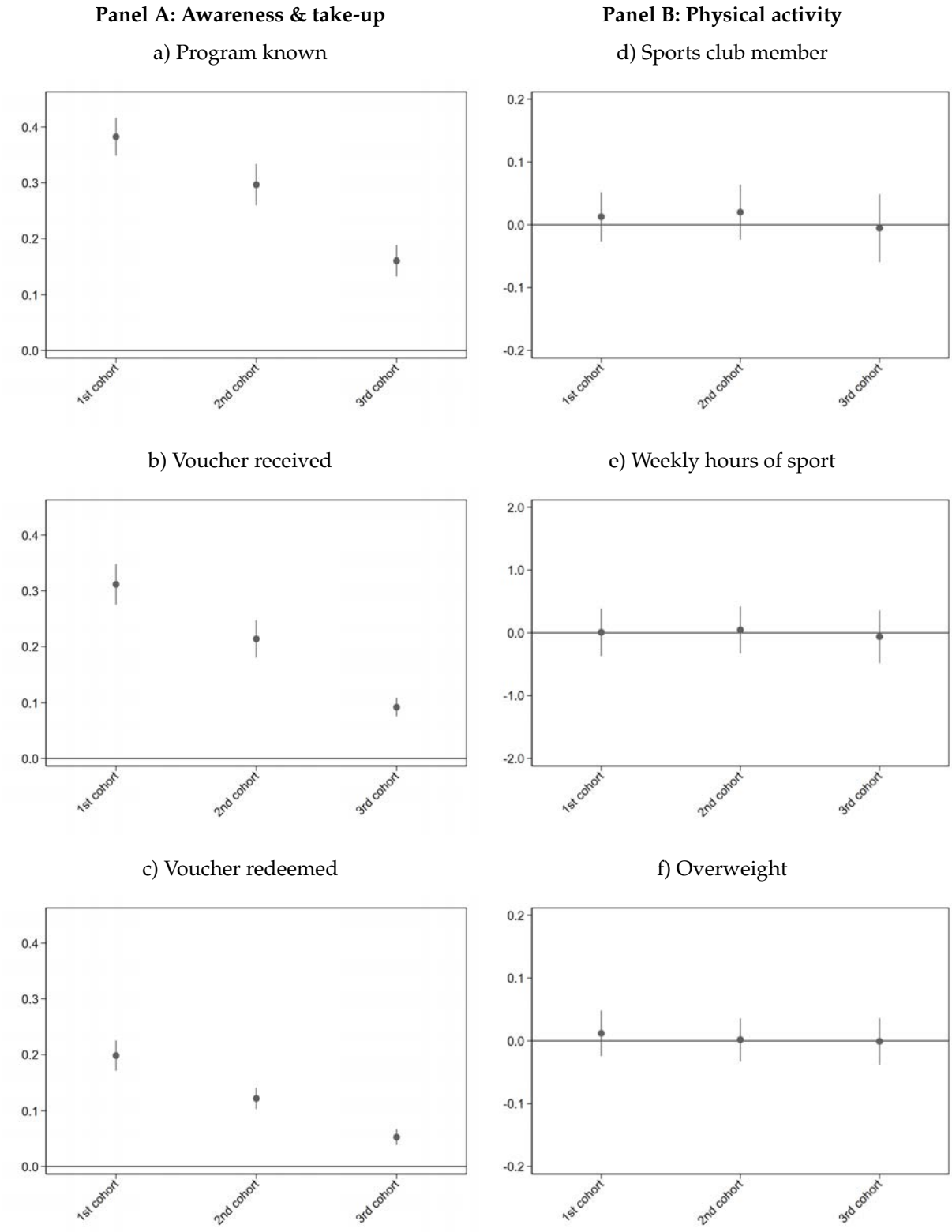
Notes: The figures display unadjusted trends of the main outcome variables by the school year in which YOLO respondents attended the third grade, before and after the start of the C2SC initiative. The treatment state is Saxony and control states are Brandenburg and Thuringia. N = 13,334.
Source: YOLO survey.

Figure 2: Effect Heterogeneity across Socio-Demographics



Notes: The figures display effect heterogeneity estimates for all six outcomes and six binary stratification variables along with 95 percent confidence bands, based on equation (3).
Source: YOLO survey.

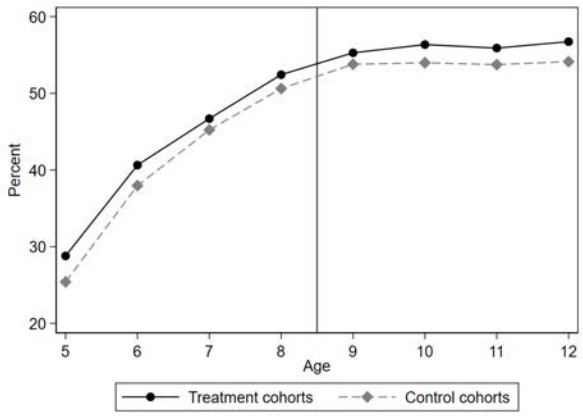
Figure 3: Effect Heterogeneity across Cohorts



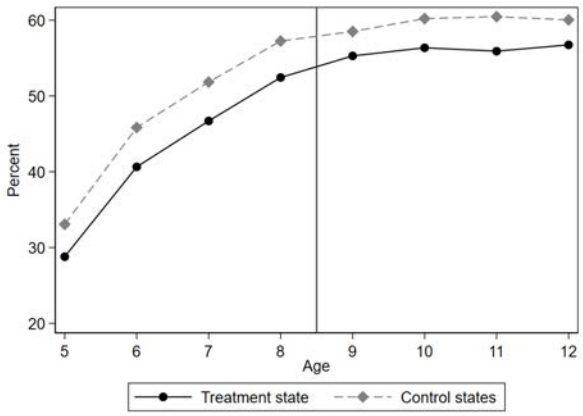
Notes: The figures display effect heterogeneity estimates for all six outcomes, by affected cohorts along with 95 percent confidence bands, based on equation (2) with municipality fixed effects.
Source: YOLO survey.

Figure 4: Sports Club Membership by Age

a) Treatment vs. control cohorts



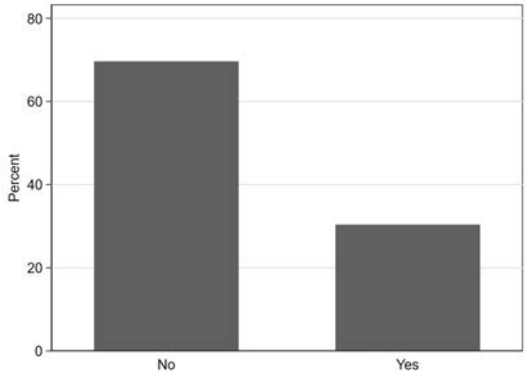
b) Treatment vs. control states



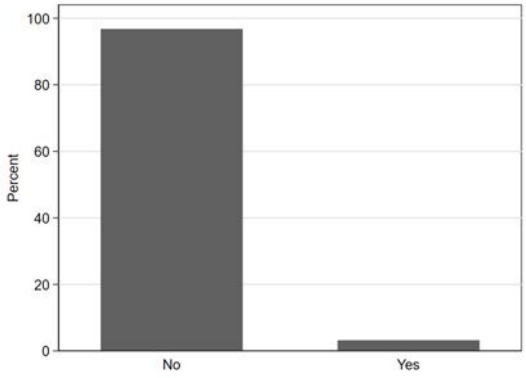
Notes: The figures display retrospectively reported sports club membership rates by child age. Figure (a) only uses data for Saxony and compares treatment cohorts (third graders in school years 2009/2010, 2010/2011, 2011/2012) to control cohorts (third graders in previous school years). Figure (b) only uses data for cohorts that attended the third grade in 2009/2010, 2010/2011, 2011/2012 and compares the treatment state (Saxony) to the control states (Brandenburg and Thuringia).
Source: YOLO survey.

Figure 5: Suggestive Evidence on Mechanisms

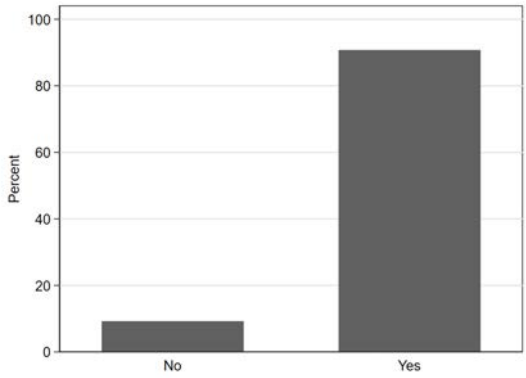
a) Tried new sport discipline(s)



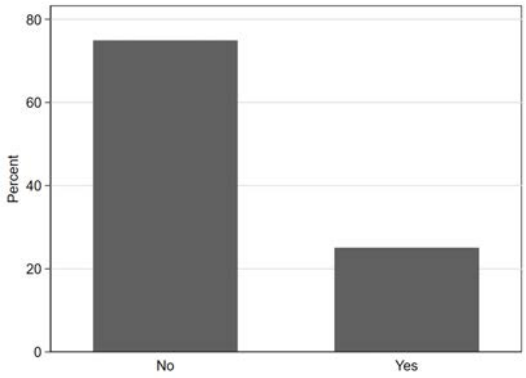
c) Could not afford membership w/o voucher



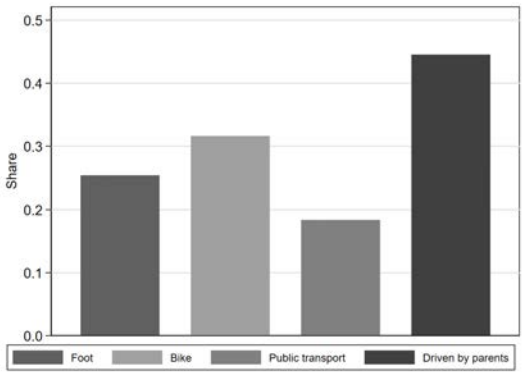
b) Could redeem voucher for desired discipline



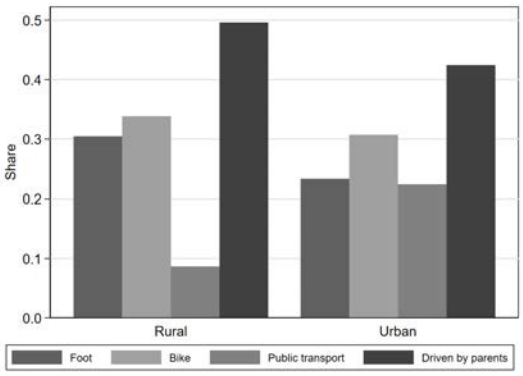
d) Parents happy to save money b/c of voucher



e) Transport as a supply-side barrier



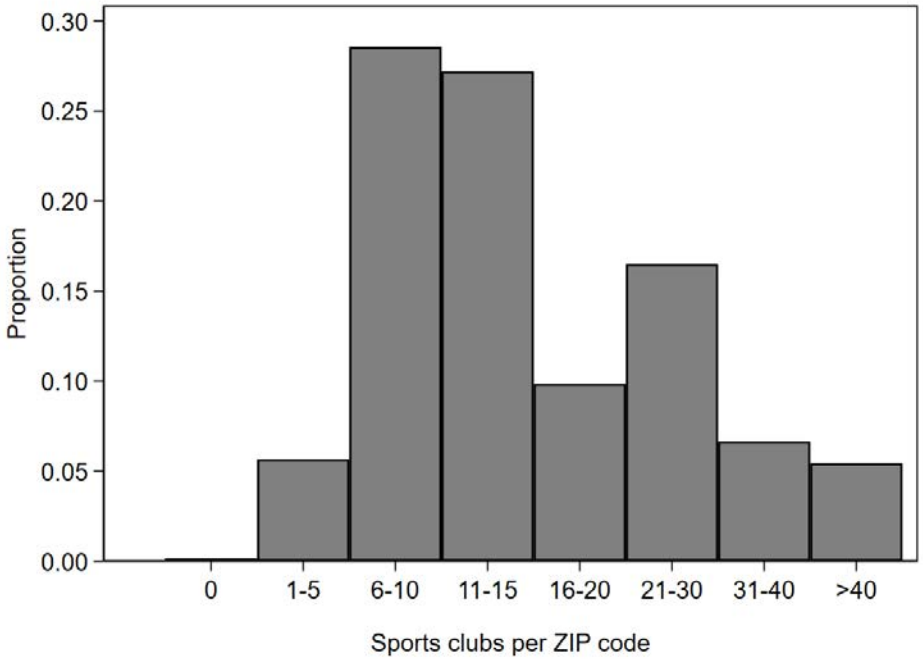
f) Mode of Transportation (Urban vs. Rural)



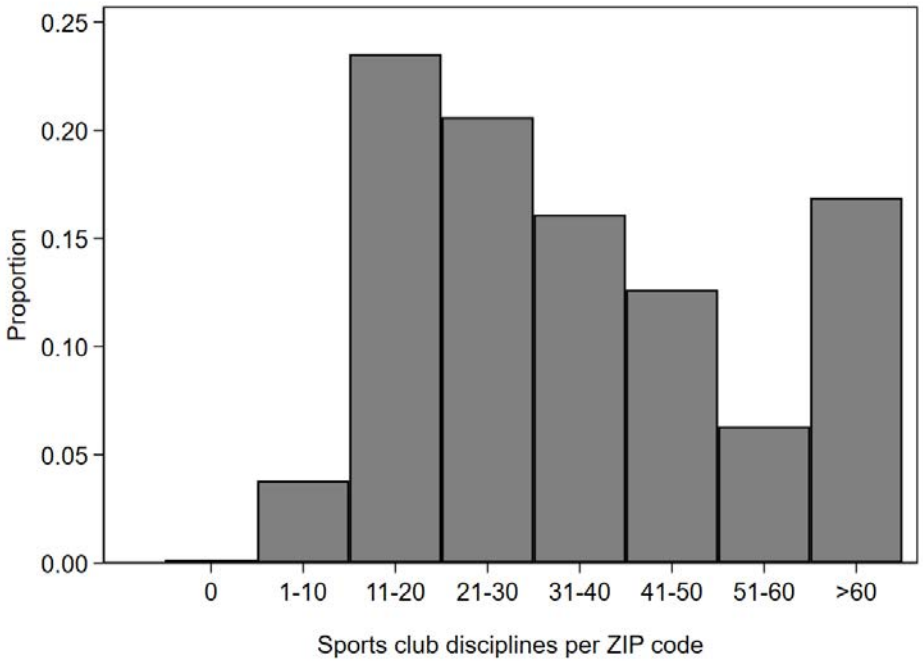
Notes: Figures a) to d) display the shares of respondents who answered the questions in the panel header with yes or no. The sample is conditional on respondents who said that they received the voucher. Figures e) and f) relate to the question "How did you get to the sports club where you redeemed your voucher?" (multiple answers possible). The sample is conditional on respondents who said that they received the voucher and remembered the mode of transport they used to get to the sports club.
Source: YOLO survey.

Figure 6: Supply-Side Restrictions? Number of Sports Clubs per ZIP Code

a) Distribution of the number of sports clubs



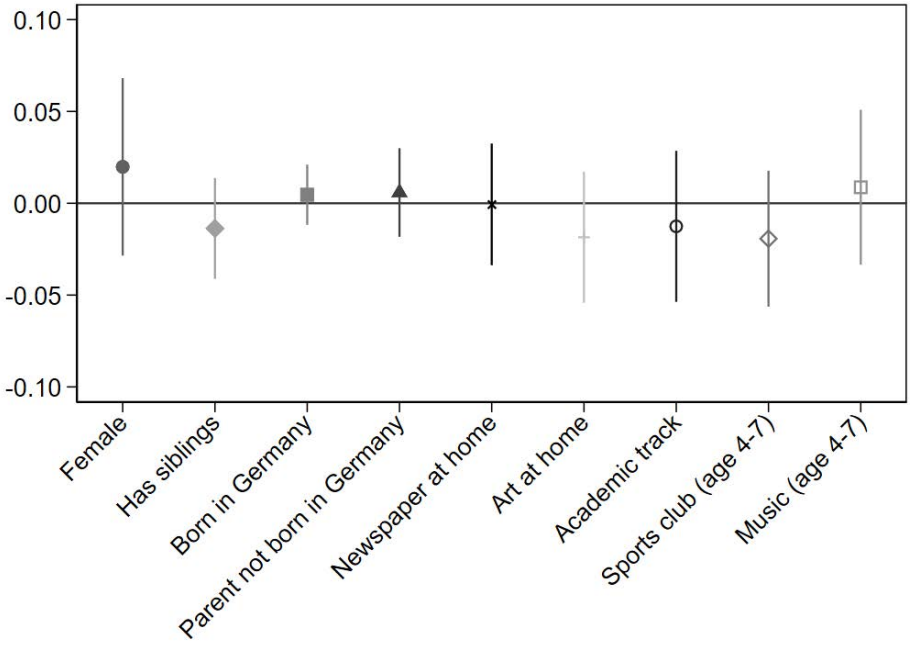
b) Distribution of the number of sports club disciplines



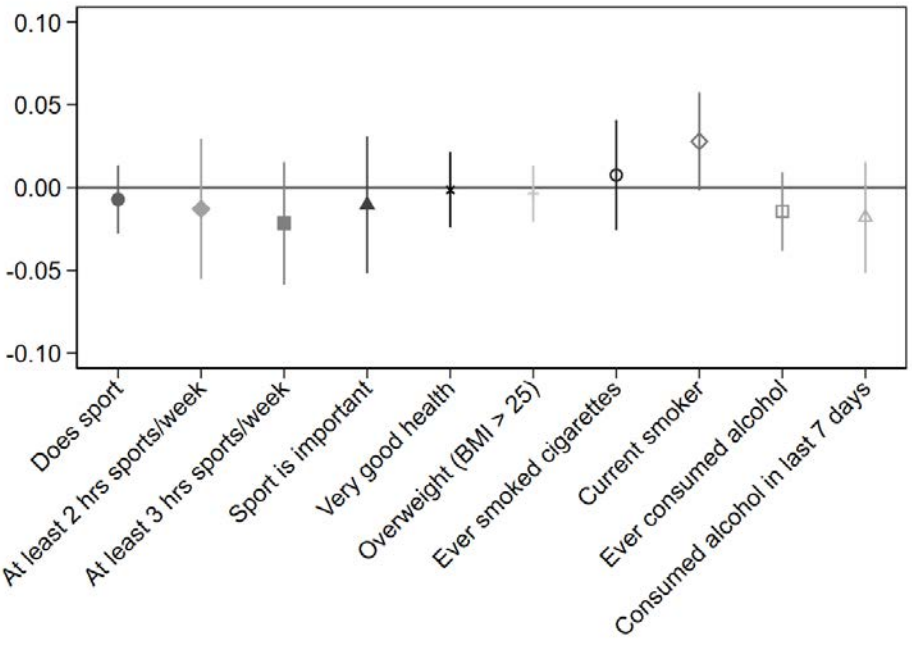
Notes: The figures display the distribution of the number of sports clubs and available disciplines across ZIP codes. Numbers are based on Saxony's 6,665 YOLO respondents from our main sample with valid ZIP code information.
Source: YOLO survey, addresses of sports clubs in Saxony.

Figure 7: Further Outcomes

a) Placebo outcomes

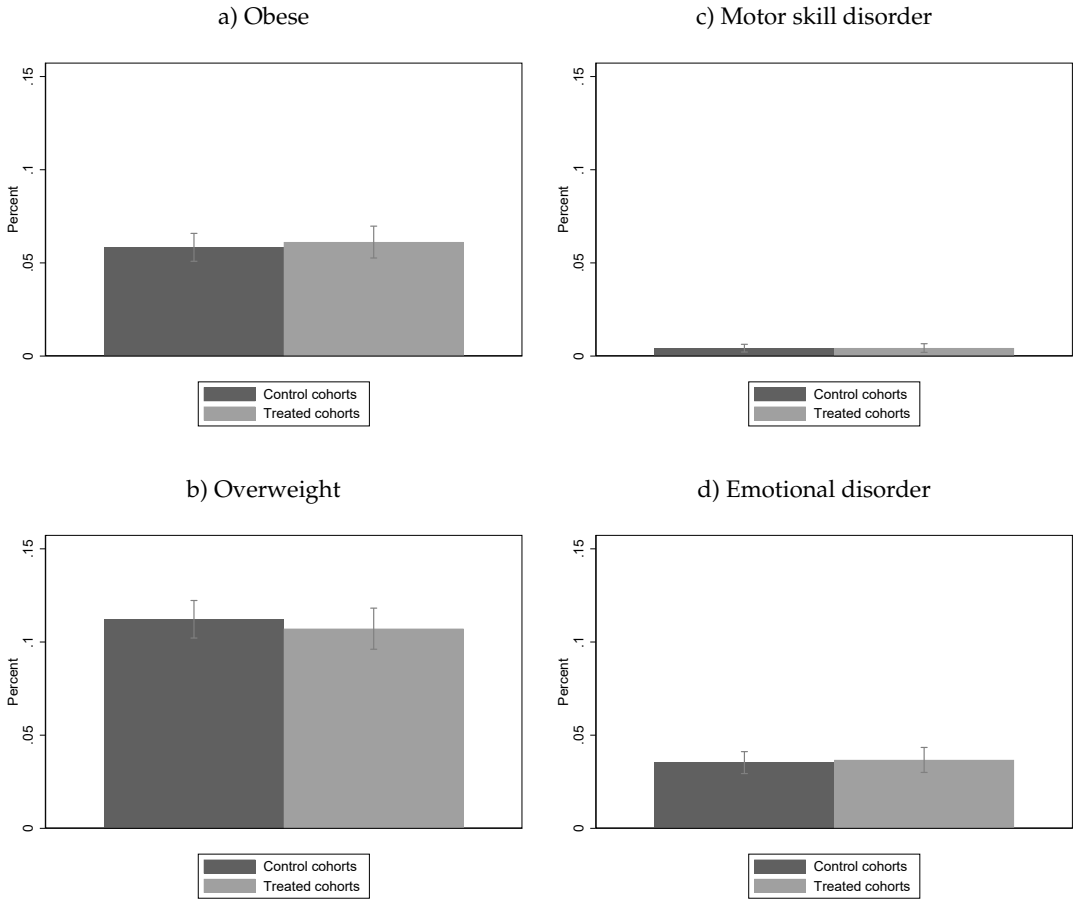


b) Additional outcomes



Notes: The figures display C2SC treatment effects on (a) placebo outcomes and (b) additional outcomes, using equation (2) with municipality fixed effects.
Source: YOLO survey.

Figure 8: Objective Health Outcomes: Treatment vs. Control Cohorts



Notes: The treated cohorts comprise those who started primary school in 2007 and 2008, received the voucher at the end of January 2010 and 2011, and had their sixth grade school medical examination between March and July of 2012 and 2013 (*Schulreihenuntersuchung*). Control cohorts comprise those who started primary school in 2004 and 2005, never received a voucher, and had their sixth grade school medical examination between March and July of 2009 and 2010. The panels display the share of students with the diagnoses indicated along with 95 percent confidence bands.
Source: School Medical Examination Data from Public Health Service (*Öffentlicher Gesundheitsdienst*).

Table 1: Summary Statistics

Variable	Mean (1)	Std. Dev. (2)	Min. (3)	Max. (4)
Background characteristics				
Age at survey	17.49	1.43	14	20
Female	0.57	0.49	0	1
City	0.61	0.49	0	1
Academic track	0.50	0.50	0	1
Newspaper at home	0.58	0.49	0	1
Art at home	0.73	0.45	0	1
Outcomes				
Program known	0.19	0.39	0	1
Voucher received	0.10	0.30	0	1
Voucher redeemed	0.06	0.24	0	1
Member of sports club	0.42	0.49	0	1
Weekly hours of sport	4.67	4.18	0	31
Overweight (BMI $>$ 25)	0.16	0.36	0	1
Treatment				
Treatment state	0.54	0.50	0	1
Treatment group	0.38	0.48	0	1

Notes: The table displays descriptive statistics for the main analysis sample. N = 13,334.
Source: YOLO survey.

Table 2: Evaluation of Sports Club Voucher Program—Main DD Results

	Base DD (1)	Twoway FE (2)	+ municip. FE (3)
Panel A: Awareness & take-up			
Program known			
Voucher	0.272***	0.272***	0.276***
	(0.014)	(0.014)	(0.014)
Voucher received			
Voucher	0.200***	0.201***	0.202***
	(0.011)	(0.011)	(0.011)
Voucher redeemed			
Voucher	0.122***	0.122***	0.122***
	(0.006)	(0.006)	(0.006)
Panel B: Physical activity and overweight			
Member of sports club			
Voucher	0.004	0.003	0.009
	(0.019)	(0.019)	(0.019)
Weekly hours of sport			
Voucher	-0.069	-0.082	-0.002
	(0.161)	(0.159)	(0.159)
Overweight			
Voucher	0.005	0.006	0.004
	(0.016)	(0.016)	(0.016)
N	13,334	13,334	13,334

Notes: The table displays the C2SC effects on various outcomes. Robust standard errors are in parentheses and clustered at the municipality level (* $p<0.1$, ** $p<0.05$, *** $p<0.01$). Each column in each panel represents one DD estimate. All regressions include state and cohort fixed effects (except (1), which only includes a dummy for the treatment group and a post-dummy). The sample includes individuals who attended the third grade between 2006 and 2010 in the German states of Saxony, Brandenburg, or Thuringia. The treatment indicator *Voucher* is one for respondents who attended third grade in Saxony in school years 2008/09, 2009/10, or 2010/11. Column (1) is based on equation (1) and column (2) on equation (2), while column (3) adds municipality fixed effects to equation (2). Column (3) is the main specification that is the basis for the subsequent analyses.
Source: YOLO survey.

Table 3: Sports Club Membership Across Child Ages

	Children		
	All (reference) (1)	Parent sample (2)	Parents' response (3)
Member of sports club at age 6			
Voucher	-0.019	0.032	-0.017
	(0.016)	(0.046)	(0.051)
Member of sports club at age 7			
Voucher	-0.025	0.014	0.016
	(0.019)	(0.044)	(0.045)
Member of sports club at age 8			
Voucher	-0.023	0.004	0.032
	(0.021)	(0.060)	(0.053)
Member of sports club at age 9			
Voucher	-0.007	0.077	0.011
	(0.018)	(0.056)	(0.049)
Member of sports club at age 10			
Voucher	0.014	0.061	-0.010
	(0.015)	(0.054)	(0.047)
Member of sports club at age 11			
Voucher	0.003	0.029	-0.023
	(0.018)	(0.058)	(0.056)
Member of sports club at age 12			
Voucher	0.008	-0.023	-0.025
	(0.016)	(0.052)	(0.057)
N	12,476	1,942	2,045

Notes: The table displays the effect of the sports club voucher program on various outcomes from the parent and child questionnaires. Column (1) uses the full sample of youth, whereas column (2) conditions on youth whose parents also responded (column [3]) to make column (2) and (3) directly comparable. Robust standard errors allowing for clustering at the municipality level are in parentheses (* $p<0.1$, ** $p<0.05$, *** $p<0.01$). Each column in each panel represents one DD estimate. Models are based on equation (2) with municipality fixed effects.
Source: YOLO survey.

Table 4: Robustness

	Temporal				Regional			Individual			Treatment		Controls		Placebo
	07-10	00-10	06-11	06-09	w/o BB	w/o TN	all	serious	sibl.	w/o older sibl.	1st grade	current	+ X_{ics}	weighted	BB vs. TN
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)	(15)
Panel A: Awareness & take-up															
Program known															
Voucher	0.245***	0.279***	0.265***	0.337***	0.271***	0.279***	0.263***	0.282***	0.279***	0.266***	0.264***	0.279***	0.287***	0.280***	-0.007
	(0.011)	(0.014)	(0.011)	(0.017)	(0.017)	(0.014)	(0.016)	(0.015)	(0.017)	(0.021)	(0.015)	(0.015)	(0.015)	(0.018)	(0.009)
Voucher received															
Voucher	0.186***	0.203***	0.196***	0.262***	0.202***	0.202***	0.196***	0.205***	0.195***	0.199***	0.196***	0.197***	0.210***	0.209***	0.000
	(0.010)	(0.011)	(0.010)	(0.017)	(0.012)	(0.011)	(0.011)	(0.013)	(0.011)	(0.014)	(0.014)	(0.012)	(0.012)	(0.013)	(0.004)
Voucher redeemed															
Voucher	0.111***	0.123***	0.118***	0.159***	0.122***	0.122***	0.117***	0.124***	0.116***	0.118***	0.116***	0.119***	0.125***	0.127***	0.001
	(0.006)	(0.006)	(0.006)	(0.010)	(0.006)	(0.006)	(0.006)	(0.006)	(0.006)	(0.007)	(0.008)	(0.006)	(0.007)	(0.008)	(0.003)
Panel B: Physical activity & overweight															
Member of sports club															
Voucher	0.011	0.006	-0.014	0.016	0.017	0.004	0.010	0.014	0.017	0.004	0.010	-0.001	0.024	0.013	0.018
	(0.019)	(0.019)	(0.015)	(0.017)	(0.024)	(0.022)	(0.020)	(0.021)	(0.020)	(0.034)	(0.017)	(0.017)	(0.019)	(0.021)	(0.028)
Weekly hours of sport															
Voucher	0.177	-0.054	-0.148	0.052	0.086	-0.034	-0.098	-0.027	0.030	-0.151	0.030	0.096	0.085	0.034	0.149
	(0.179)	(0.164)	(0.119)	(0.162)	(0.185)	(0.189)	(0.154)	(0.174)	(0.169)	(0.205)	(0.130)	(0.139)	(0.150)	(0.184)	(0.204)
Overweight															
Voucher	0.014	0.010	0.005	0.006	0.018	-0.004	0.000	0.005	-0.002	0.031	0.006	0.005	-0.001	-0.003	0.022
	(0.022)	(0.015)	(0.010)	(0.016)	(0.022)	(0.018)	(0.015)	(0.019)	(0.016)	(0.029)	(0.014)	(0.016)	(0.017)	(0.016)	(0.024)
N	11,686	13,506	16,082	10,044	9,572	10,973	14,720	10,836	12,481	5,891	13,421	13,859	12,114	12,004	6,123

Notes: The table tests the robustness of the C2SC effects for all main outcomes. Robust standard errors allowing for clustering at the municipality level are in parentheses (* $p<0.1$, ** $p<0.05$, *** $p<0.01$). All regressions include state and cohort fixed effects. The main specification includes individuals who attended the third grade in the years 2006-2010 in Saxony, Brandenburg (BB), or Thuringia (TN). The treatment indicator is based on the year and state when an individual attended the third grade (except in columns [11] and [12]). The column headers indicate the type of robustness check, see main text for details. "Sibl." stands for "sibling." Column (8)—"serious"—conditions on respondents who spent at least 10 minutes on the survey and whose registry and self-reported information match exactly. Column (9) excludes respondents from the pre-treatment cohorts whose siblings also received a YOLO invitation. Column (10) excludes youth with an older sibling.
Source: YOLO survey.

Table 5: Using Parents' Responses for Program-Related Outcomes

	Children		
	All (reference) (1)	Parent sample (2)	Parents' response (3)
Program known			
Voucher	0.276***	0.322***	0.271***
	(0.014)	(0.017)	(0.024)
Voucher received			
Voucher	0.202***	0.233***	0.275***
	(0.011)	(0.017)	(0.025)
Voucher redeemed			
Voucher	0.122***	0.144***	0.167***
	(0.006)	(0.011)	(0.021)
N	13,334	2,045	2,045

Notes: The table displays the effect of the sports club voucher program on our main outcomes using information from the parent and child questionnaires. Column (1) uses the full sample of youth, whereas column (2) conditions on youth whose parents also responded (column [3]) to make column (2) and (3) directly comparable. Robust standard errors allowing for clustering at the municipality level are in parentheses (* $p<0.1$, ** $p<0.05$, *** $p<0.01$). Models are based on equation (2) with municipality fixed effects.
Source: YOLO survey.

Table 6: Synthetic Control Group Results

	Synthetic control	
	Version 1	Version 2
	(1)	(2)
Panel A: Awareness & take-up		
Program known		
Voucher	0.285***	0.280***
	(0.017)	(0.017)
Voucher received		
Voucher	0.207***	0.203***
	(0.014)	(0.014)
Voucher redeemed		
Voucher	0.127***	0.123***
	(0.007)	(0.008)
Panel B: Physical activity and overweight		
Member of sports club		
Voucher	-0.008	-0.016
	(0.022)	(0.021)
Weekly hours of sport		
Voucher	-0.048	-0.206
	(0.202)	(0.290)
Overweight		
Voucher	-0.017	-0.012
	(0.023)	(0.021)
$N_{\texttt{cohort}-\texttt{municipality}}$	435	435

Notes: The table displays the effects of the sports club voucher program on our main outcomes, based on municipality-level regressions and a synthetic control group approach. The synthetic control group is based on "entropy balancing" (Hainmueller, 2012) and reweights the municipalities in the control states such that the two pre-treatment cohorts have the same means than the treatment state with respect to specific variables. Version 1 considers only *member of sports club*, *weekly hours of sport*, and *overweight* for the construction of the synthetic control group, while version 2 additionally relies on all pre-determined grouping variables used in the heterogeneity analyses of Figure 2. All regressions include municipality and cohort fixed effects. The estimations use a balanced sample of 87 municipalities that had survey respondents in all five cohorts. *N* refers to the number of cohort-municipality observations. Robust standard errors are in parentheses and clustered at the municipality level (* $p<0.1$, ** $p<0.05$, *** $p<0.01$).
Source: YOLO survey.

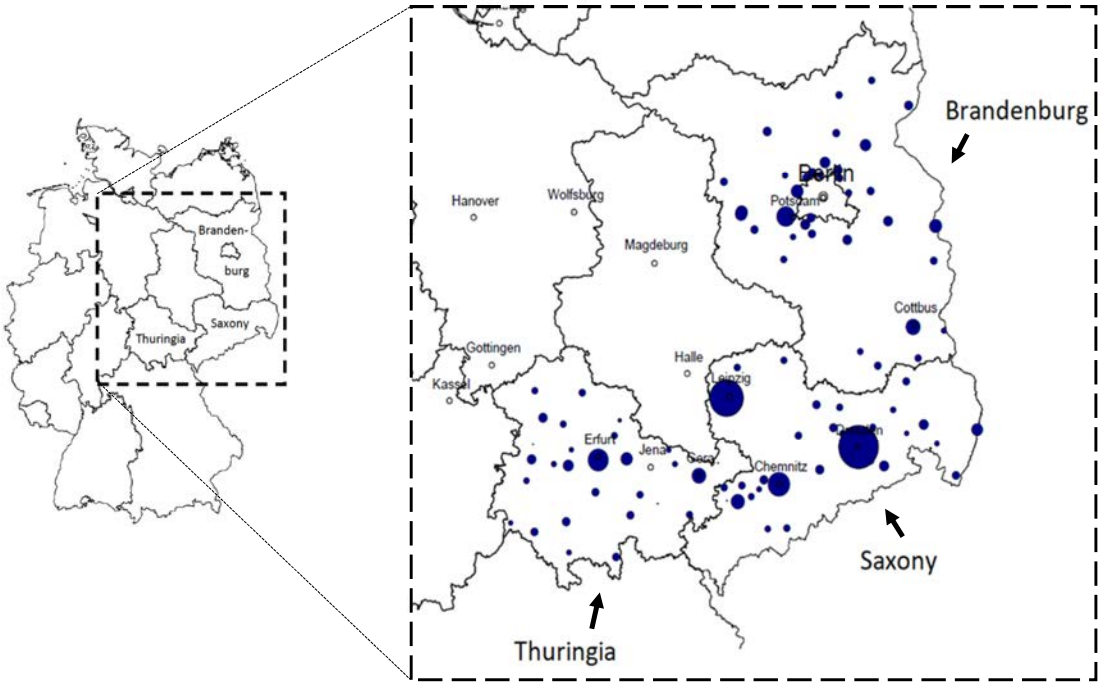
Appendix A: YOLO Survey

Figure A1: Example of a Sports Club Voucher

a) German version (original)

b) English version (translation)

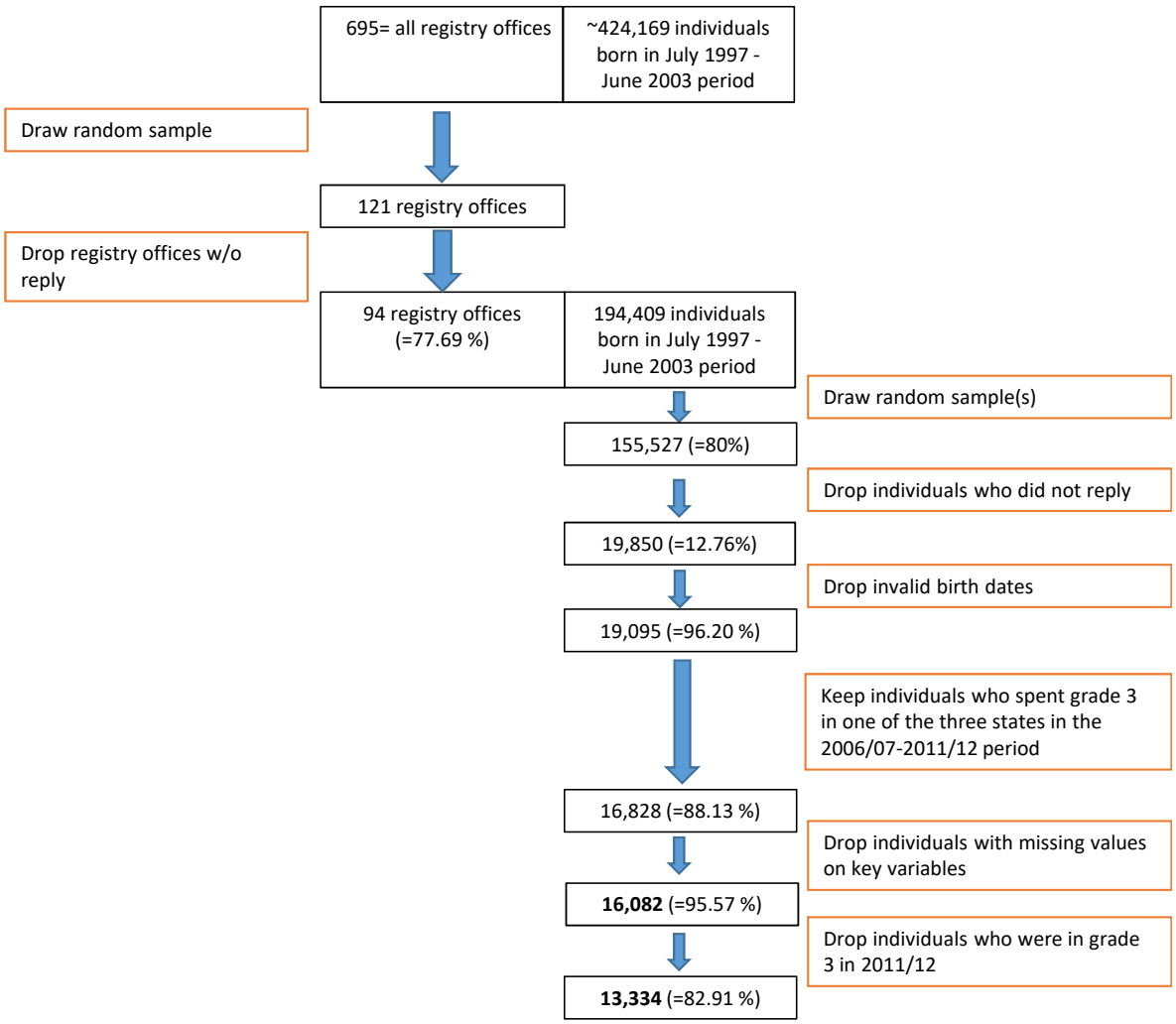
Figure A2: Map of Resident Registries



Notes: The map on the left shows all 16 German federal states. The map on the right displays the participating registries in the treatment (Saxony) and control states (Brandenburg, Thuringia). The circles are proportional to population size. *Source:* own illustration.

Sampling and Sample Sizes

Figure A3: YOLO Sampling and Sample Size



Sampling Procedure. At the beginning of 2018, a total of 424,169 individuals who were born between July 1997 and June 2002 lived in the three states of Brandenburg, Saxony, and Thuringia (Statistisches Bundesamt, 2020).[36] There were 695 registry offices in these three states.

In the first sampling stage, we randomly selected 121 out of these total of 695 registry offices (with sampling probabilities proportional to population size). About 78 percent or 94 registry offices replied. These 94 registry offices cover about 46 percent of the total target population in the three federal states (194,409 youth). Note that because we sampled registry offices proportional to population size, the share of sampled individuals is much larger than the share of sampled registry offices. Due to German data privacy regulations, we were only allowed to ask for an 80 percent random sample of all households in our target group from these 94 registry offices, as shown in Figure A2. Hence we obtained 155,527 (physical) addresses for youth in the target cohorts.

In the next step, we mailed a research invitation letter (see Figure A1) to each of these 155,527 youth and invited them to participate in our YOLO survey online. A total of 19,850 individuals replied and completed the online survey (12.76 percent). After discarding individuals with self-reported birth dates outside our target cohorts, we were left with 19,095 respondents. Of these, we considered only those who reported that they attended third grade in one of the three states. This reduced the sample size by about 12 percent to 16,828 individuals. Additionally, we only kept respondents who answered the relevant questions to construct our six main outcome variables (16,082 individuals).

Moreover, in our main analysis, we disregard third graders in school year 2011/12, which further reduces the sample size to 13,334 individuals. We discard this cohort because i) the Educational Package (*Bildungs- und Teilhabepaket*) that covered sports club membership fees for welfare recipients came into effect on April 1, 2011 and therefore affected this cohort (see Section 2.2) and ii) because this is the first cohort that did not receive vouchers and disappointment effects could therefore arise.

[36] Note that this number is only an approximation as Statistisches Bundesamt (2020) only provides the numbers based on birth years.

Figure A4: Information Letter for Survey Respondents

a) German version (original)

Universität Hamburg
DER FORSCHUNG | DER LEHRE | DER BILDUNG

FAKULTÄT
FÜR WIRTSCHAFTS- UND
SOZIALWISSENSCHAFTEN

UHH - WiSO Fakultät - Professur VWL, insb. Mikroökonometrie
Esplanade 36 - D-20354 Hamburg

Prof. Thomas Siedler (PhD)

Universität Hamburg
Fakultät für Wirtschafts- und Sozialwissenschaften
Fachbereich Volkswirtschaftslehre
Professur für Volkswirtschaftslehre, insb. Mikroökonometrie
www.wiso.uni-hamburg.de/fachbereich-vwl/professuren/siedler.html
E-Mail: umfrage-mikro@wiso.uni-hamburg.de
Tel.: 040 42838-9459

19.06.2018

Teilnahme an einer wissenschaftlichen Umfrage

Sehr geehrte Frau X,

wir sind Wissenschaftler an der Universität Hamburg und führen eine Online-Umfrage zum Freizeitverhalten von Jugendlichen und jungen Erwachsenen durch. Hiermit laden wir Sie herzlich ein, an dieser Umfrage teilzunehmen.[1] Unter allen Teilnehmenden verlosen wir zwei aktuelle iPads im Wert von je 500 Euro und zehn Amazon-Gutscheine im Wert von je 20 Euro. Die Teilnahme an der Umfrage ist freiwillig. Den Zugang zur Umfrage sowie weitere Informationen erhalten Sie mit folgenden Zugangsdaten:

Link für Sie:	www.umfrage-uhh.de/uc/u
Passwort:	XXYYZZ
Link für Eltern:	www.umfrage-uhh.de/uc/eltern

Bitte geben Sie diese Zugangsdaten auch an einen Elternteil weiter. Ihre Mutter oder Ihr Vater kann unter dem zuletzt angegebenen Link mit dem gleichen Passwort wie Sie ebenfalls an der Umfrage teilnehmen.

Wir bitten Sie, bis spätestens zwei Wochen nach Erhalt dieses Briefs teilzunehmen. Herzlichen Dank!

Mit freundlichen Grüßen

...

[1] Ihre Anschrift haben wir auf Grundlage des § 46 Bundesmeldegesetz von der für Sie zuständigen Meldebehörde erhalten. Ihr Name und Ihre Anschrift werden aus unserer Datenbank gelöscht, nachdem wir Sie angeschrieben haben.

b) English version (translation)

Universität Hamburg
DER FORSCHUNG | DER LEHRE | DER BILDUNG

FAKULTÄT
FÜR WIRTSCHAFTS- UND
SOZIALWISSENSCHAFTEN

UHH - WiSO Fakultät - Professur VWL, insb. Mikroökonometrie
Esplanade 36 - D-20354 Hamburg

Prof. Thomas Siedler (PhD)

Universität Hamburg
Fakultät für Wirtschafts- und Sozialwissenschaften
Fachbereich Volkswirtschaftslehre
Professur für Volkswirtschaftslehre, insb. Mikroökonometrie
www.wiso.uni-hamburg.de/fachbereich-vwl/professuren/siedler.html
E-Mail: umfrage-mikro@wiso.uni-hamburg.de
Tel.: 040 42838-9459

June 19, 2018

Participation in a scientific survey

Dear Sir/Madam,

We are researchers at the University of Hamburg conducting an online survey on the leisure behavior of adolescents and young adults. We cordially invite you to take part in this survey.[1] As a thank you for participating your name will be entered into a lottery, where you will have the chance to win two new iPads worth €500 each and ten Amazon vouchers worth €20 each. Participation in the survey is voluntary. You can access the survey and obtain further information using the following access data:

Link for you:	www.umfrage-uhh.de/uc/u
Password:	XXYYZZ
Link for your parents:	www.umfrage-uhh.de/uc/eltern

Please pass this access data on to one parent. This parent can also take part in the survey using the same password as you via the second link listed above.

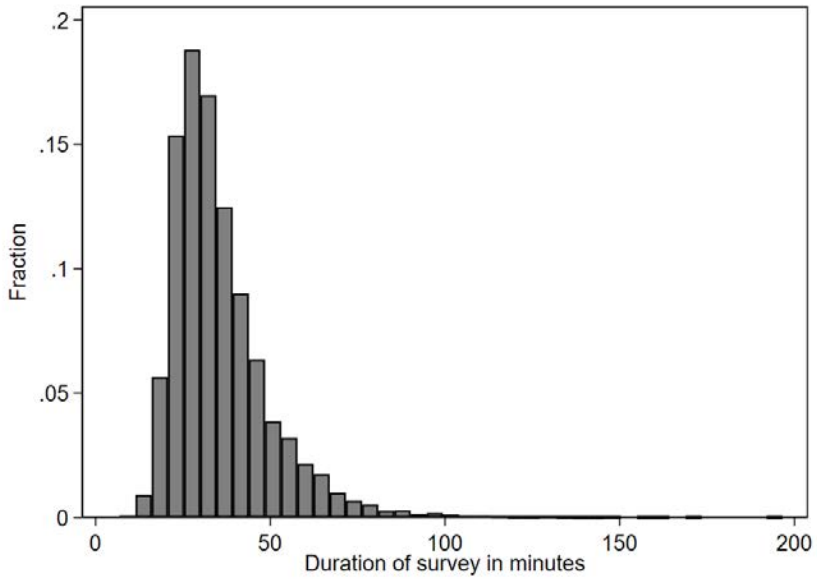
We would request that you participate within two weeks of receiving this letter at the latest. Many thanks for your help!

Yours sincerely

...

[1] We obtained your address from the registration authority where you are registered in accordance with Section 46 of the Federal Registration Act (*Bundesmeldegesetz*). Your name and address will be deleted from our database after we have written to you.

Figure A5: Duration of Survey in Minutes



Notes: The figure displays the minutes needed to complete the YOLO survey for the individuals in the main analysis sample.
Source: YOLO survey.

Table A1: Stylized School Cohort Development

Birth date	School enrollment (1)	In 3rd grade (2)	C2SC voucher (3)
July 1997–June 1998	2004	2006/2007	No
July 1998–June 1999	2005	2007/2008	No
July 1999–June 2000	2006	2008/2009	Yes
July 2000–June 2001	2007	2009/2010	Yes
July 2001–June 2002	2008	2010/2011	Yes

Notes: The table displays the stylized relationship between birth cohorts and school cohorts as well as their eligibility status for the C2SC voucher. The relationship is stylized in the sense that it does not consider deviations from this path (e.g., red-shirting, grade repetitions), which are, however, incorporated in the construction of the main treatment indicator.

Table A2: Summary Statistics: Treatment vs. Control States

Variable	Treatment state (1)	Control states (2)	Norm.diff. (3)
Background characteristics			
Female	0.57	0.57	0.01
Has siblings	0.87	0.84	0.06
Born in Germany	0.96	0.97	-0.02
Parent not born in Germany	0.14	0.12	0.04
Newspaper at home	0.54	0.63	-0.12
Art at home	0.73	0.72	0.02
Academic track	0.49	0.50	-0.02
Sports club age 4-7	0.52	0.56	-0.06
Music lessons age 4-7	0.47	0.45	0.04
$\geq$ 1hr sports per week	0.89	0.89	0.00
$\geq$ 2hrs sports per week	0.76	0.77	-0.02
$\geq$ 3hrs sports per week	0.62	0.64	-0.03
Sport is important	0.57	0.57	0.00
Very good health	0.26	0.28	-0.04
Obese (BMI>30)	0.04	0.04	-0.02
Ever smoked cigarettes	0.53	0.57	-0.06
Current smoker	0.20	0.22	-0.02
Ever consumed alcohol	0.81	0.84	-0.05
Alcohol in last 7 days	0.51	0.52	-0.02
Age at survey	17.56	17.42	0.07
City	0.76	0.44	0.50
Outcomes			
Program known	0.32	0.03	0.60
Voucher received	0.18	0.00	0.45
Voucher redeemed	0.11	0.00	0.34
Member of sports club	0.41	0.43	-0.02
Weekly hours of sport	4.63	4.70	-0.01
Overweight (BMI>25)	0.15	0.17	-0.03

Notes: The table displays descriptive statistics for the main analysis sample, separately for Saxony (treatment state) and Brandenburg and Thuringia (control states). Norm. diff. stands for the "normalized difference", which is defined for each variable x as $ND_x = (\overline{x_1} - \overline{x_0}) / \sqrt{(s^2_{x1} + s^2_{x0})}$, where $\overline{x_1}$ and $\overline{x_0}$ are the sample means of the two groups and s^2_{x1} and s^2_{x0} the corresponding variances. According to Imbens and Wooldridge (2009), a normalized difference of more than 0.25 indicates substantial covariate imbalance.
Source: YOLO survey.

Table A3: Registry vs. Self-Reported Socio-Demographics

	Percentage survey = registry	N
	(1)	(2)
Female	0.995	13,331
German nationality	0.995	13,040
Year of birth	0.993	12,105
Month and year of birth	0.985	12,105
Day, month, and year of birth	0.974	12,105

Notes: The table displays the share of individuals in our main sample for which the registry information matches the self-reported information.
Source: YOLO survey and registry information.

Table A4: Administrative Data: YOLO Participants vs. Non-Participants

Variable	YOLO-participants (1)	Non-participants (2)	Norm. diff. (3)
Female	0.56	0.47	0.13
German nationality	0.96	0.90	0.15
Saxony	0.55	0.51	0.05
Year of birth 1997	0.06	0.09	-0.08
Year of birth 1998	0.12	0.18	-0.11
Year of birth 1999	0.15	0.17	-0.05
Year of birth 2000	0.18	0.16	0.03
Year of birth 2001	0.19	0.15	0.07
Year of birth 2002	0.19	0.15	0.08
Year of birth 2003	0.08	0.07	0.03

Notes: The table compares YOLO participants with non-participants based on registry information.
Source: Registry information.

Table A5: Comparison of YOLO and SOEP Participants

Variable	YOLO (1)	SOEP (2)	Norm. diff. (3)
Socio-demographic variables			
Female	0.57	0.54	0.04
German citizenship	0.96	0.98	-0.06
Born in Germany	0.95	0.97	-0.07
Has siblings	0.86	0.86	0.00
Still in school	0.84	0.94	-0.24
Leisure time activities			
Does sport	0.72	0.74	-0.03
Does sport in a club	0.31	0.31	0.00
Involved in music	0.33	0.32	0.01
Music lessons outside school	0.25	0.19	0.11
Watches TV, videos	1.39	1.37	0.02
Plays computer games	2.46	2.59	-0.06
Listens to music	1.21	1.15	0.08
Plays music, sings	3.65	3.88	-0.11
Does sport	2.30	2.38	-0.05
Dances or acts	4.17	4.02	0.09
Reads	2.68	2.76	-0.05
Does volunteer work	4.35	4.49	-0.10
Does nothing	2.32	2.02	0.18
Best friend	2.25	2.07	0.14
Youth/recreation centre	4.76	4.38	0.30
Church/religious events	4.59	4.48	0.09
Personality traits and attitudes			
Risk attitude	5.56	5.72	-0.05
Internal locus-of-control	0.00	0.11	-0.07
External locus-of-control	-0.03	0.05	-0.06
Works carefully	5.65	5.16	0.27
Communicative	5.04	5.12	-0.04
Abrasive towards others	3.40	3.22	0.08
Introduces new ideas	4.81	4.73	0.04
Often worries	5.08	5.00	0.03
Can forgive others	5.70	5.60	0.06
Is lazy	4.20	4.07	0.05
Is outgoing/sociable	4.77	4.97	-0.09
Importance of aesthetics	4.76	4.42	0.13
Is nervous	4.32	4.24	0.03
Carries out duties efficiently	5.49	5.07	0.25
Is reserved	4.23	4.32	-0.04
Is considerate, friendly	6.00	5.95	0.03
Has a lively imagination	5.35	5.24	0.05
Is relaxed/unstressed	4.48	4.15	0.15
Is curious	5.50	5.21	0.16
Is positive about oneself	4.97	4.90	0.03

Notes: The table compares YOLO participants with SOEP participants. We use the SOEP youth questionnaire and the year respondents turned 17. To make the two samples comparable, both are restricted to individuals born between July 1997 and July 2000 (2000 is the last available cohort in the SOEP youth questionnaire and July 1997 is the first cohort in YOLO), who live in Saxony, Brandenburg, or Thuringia. SOEP observations are weighted with SOEP weights. Imbens and Wooldridge (2009) provide details on how to calculate the normalized difference. The construction of the locus of control variables follows Peter and Spiess (2016) for the pooled SOEP-YOLO sample. *Source*: YOLO survey and SOEP.

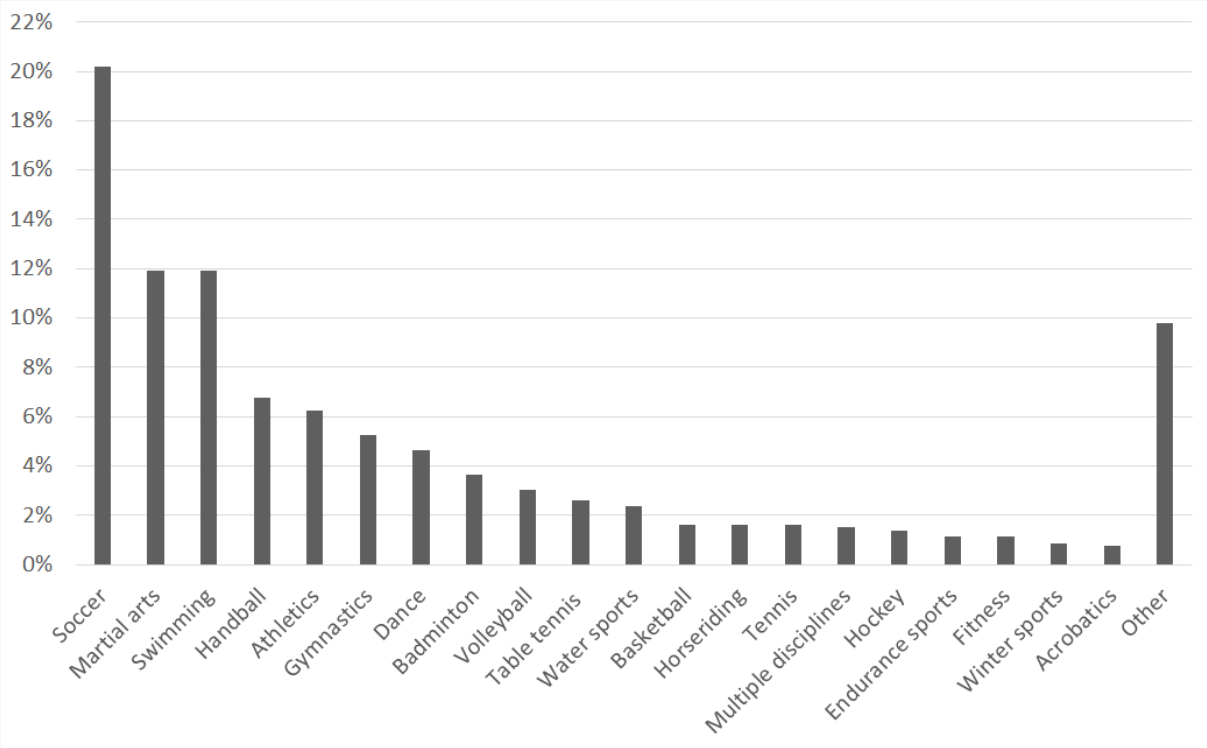
Table A6: Survey Participation as Outcome in DD Framework

	Cohorts 1997-2003 (1)	Cohorts 1997-2002 (2)
Voucher	0.004	0.007
	(0.004)	(0.006)
N	141,758	120,528

Notes: The table displays the "effect" of the voucher program on participation in the YOLO survey based on our DD framework in equation (2) and registry information on state and date of birth. The treatment indicator is assigned based on birth dates. The binary outcome takes on the value one if an individual participated in the survey and zero otherwise.
Source: Registry information.

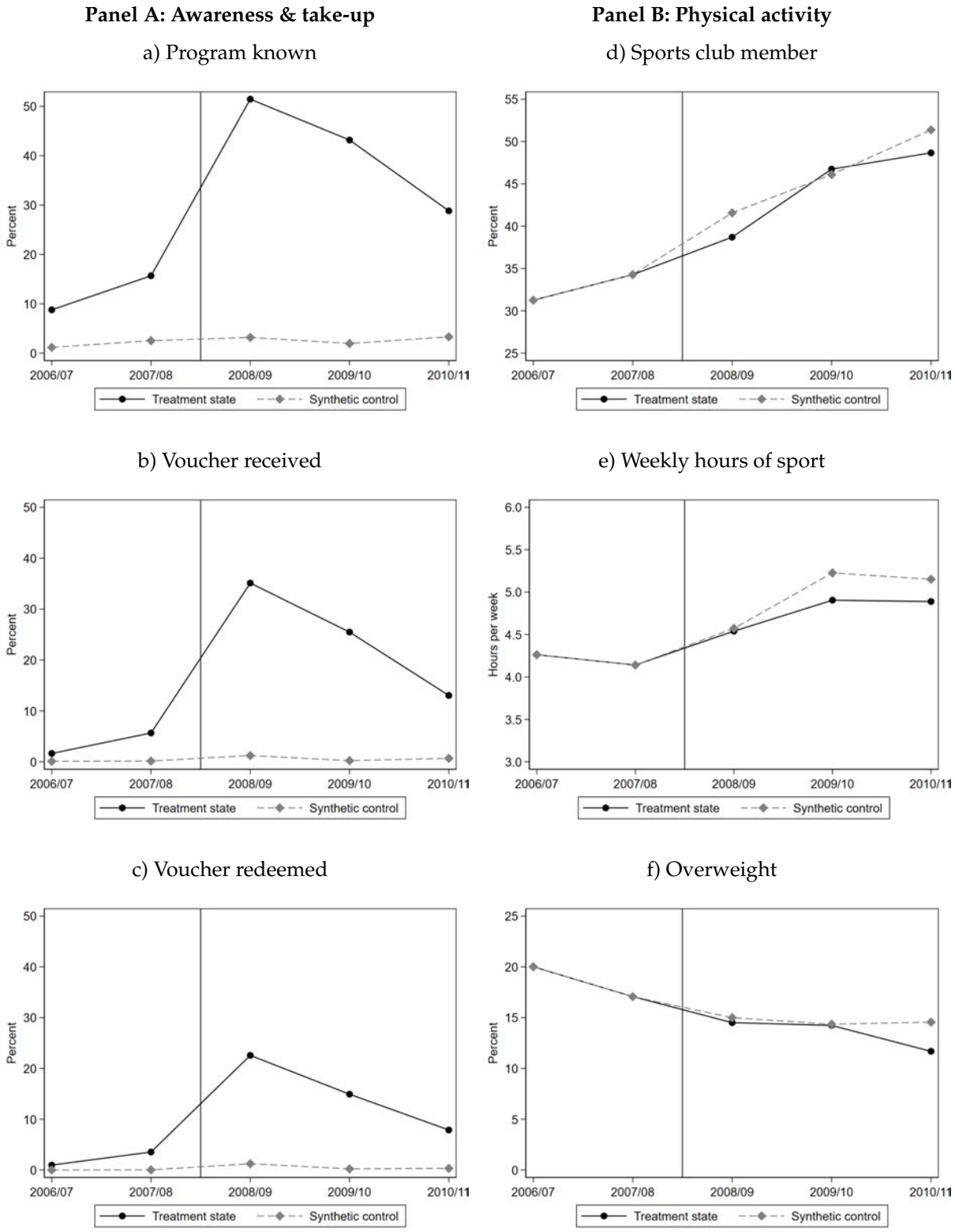
Appendix B: Empirical Results YOLO

Figure B1: Sports Disciplines for which Vouchers Were Redeemed



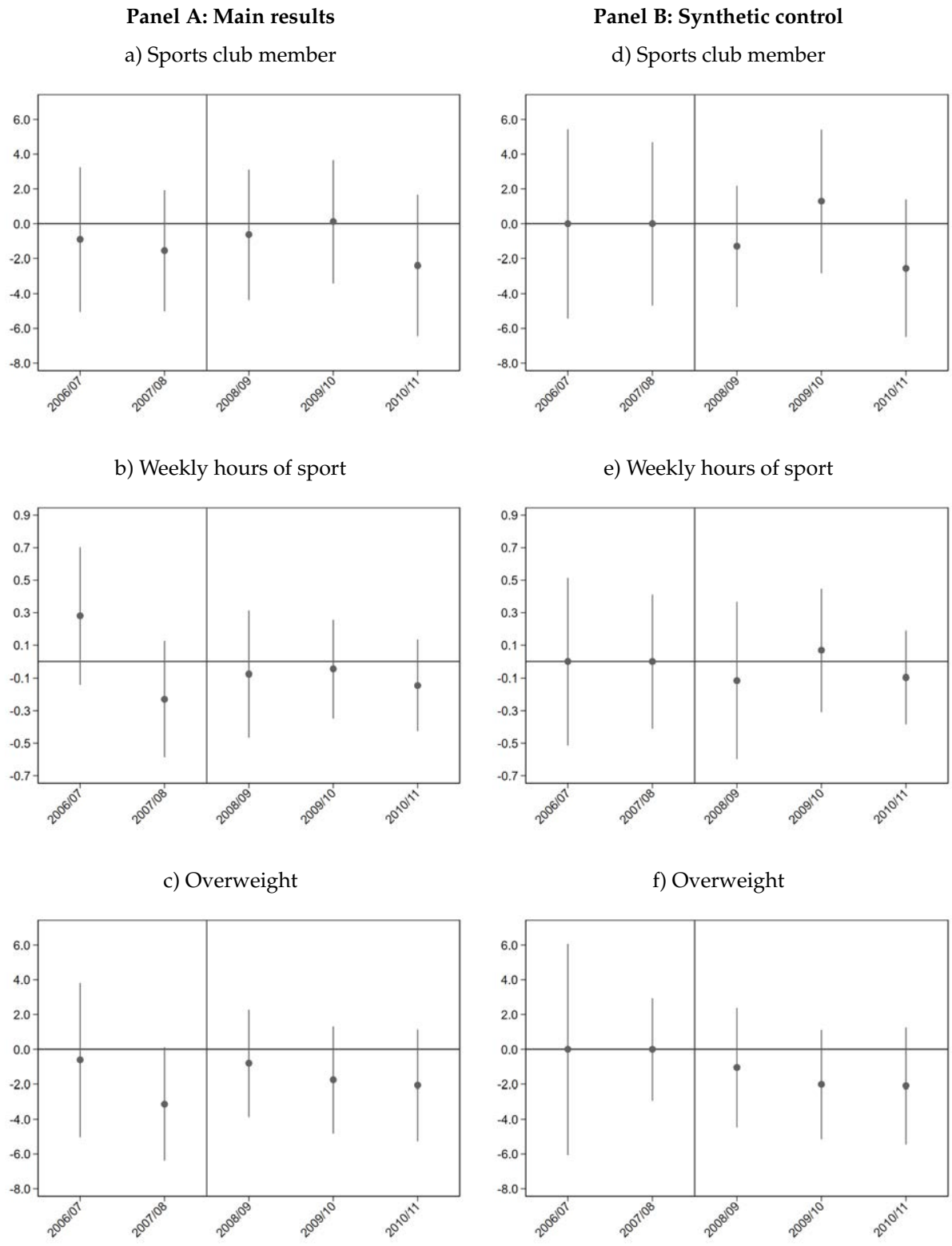
Notes: The figure displays the sports disciplines for which vouchers were redeemed. The sample includes only respondents who answered that they redeemed the voucher (N=798).
Source: YOLO survey.

Figure B2: Development of Outcome Variables—Synthetic Control Group



Notes: The figures show the main outcome variables by the school year during which YOLO respondents attended the third grade, before and after the start of the C2SC initiative. The figures compares the treatment state of Saxony to a synthetic control group based on municipalities in Brandenburg and Thuringia. See the notes to Table 6 for further details on the construction of the synthetic control group. *Source*: YOLO survey.

Figure B3: Outcome Difference—Treatment vs. Control States



Notes: While Figures 1 and B2 display the mean of the main outcome variables for the treatment and control states, this figure displays the *differences* between the treatment and control states based on the main sample (Panel A) and on the synthetic control group (Panel B). Note that these are simple averages and not adjusted for covariates.
Source: YOLO survey.

Table B1: Difference-in-Differences: Heterogeneity

	Sports club age 4-7		Newspaper		Art at home		Academic track		Female		Urban	
	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
Panel A: Awareness & take-up												
Program known												
Voucher	0.254***	0.293***	0.277***	0.271***	0.247***	0.299***	0.192***	0.306***	0.201***	0.283***	0.256***	0.297***
	(0.025)	(0.012)	(0.025)	(0.017)	(0.014)	(0.018)	(0.013)	(0.017)	(0.019)	(0.010)	(0.017)	(0.016)
Voucher received												
Voucher	0.193***	0.210***	0.196***	0.206***	0.188***	0.211***	0.140***	0.223***	0.156***	0.215***	0.183***	0.227***
	(0.016)	(0.010)	(0.020)	(0.014)	(0.014)	(0.012)	(0.007)	(0.014)	(0.013)	(0.012)	(0.009)	(0.017)
Voucher redeemed												
Voucher	0.118***	0.126***	0.130***	0.120***	0.106***	0.133***	0.080***	0.135***	0.092***	0.132***	0.085***	0.156***
	(0.014)	(0.009)	(0.016)	(0.007)	(0.009)	(0.007)	(0.009)	(0.010)	(0.007)	(0.009)	(0.007)	(0.011)
Panel B: Physical activity												
Member of sports club												
Voucher	0.015	0.006	0.009	0.008	-0.001	0.021	-0.002	0.011	0.011	0.013	0.027	0.019
	(0.026)	(0.021)	(0.042)	(0.024)	(0.029)	(0.021)	(0.039)	(0.019)	(0.024)	(0.030)	(0.022)	(0.029)
Weekly hours of sport												
Voucher	0.328	-0.218	0.065	-0.118	-0.292	0.198	-0.099	-0.001	-0.126	0.263	0.028	0.148
	(0.223)	(0.201)	(0.272)	(0.222)	(0.242)	(0.172)	(0.242)	(0.168)	(0.207)	(0.258)	(0.201)	(0.209)
Overweight												
Voucher	0.006	0.001	0.015	0.004	0.017	-0.004	-0.006	0.011	-0.018	0.012	-0.015	0.007
	(0.025)	(0.020)	(0.026)	(0.021)	(0.019)	(0.022)	(0.036)	(0.018)	(0.020)	(0.025)	(0.021)	(0.024)
N	5,722	7,609	5,192	8,142	5,576	7,666	3,612	9,643	6,661	6,601	5,717	6,759

Notes: The table displays the effect of the C2SC initiative for various subgroups as indicated by the column header. Robust standard errors allowing for clustering at the municipality level are in parentheses (* $p<0.1$, ** $p<0.05$, *** $p<0.01$). Each column in each panel represents one DD estimate based on the subsample as indicated by the column header. Models are based on equation (2) with municipality fixed effects.
Source: YOLO survey.

Table B2: Alternative Aged-Based Difference-in-Differences Models

	Within Saxony estimation		Within treated cohort estimation	
	DD	Event	DD	Event
	(1)	(2)	(3)	(4)
Overall effect	-0.002		0.011	
	(0.009)		(0.008)	
(Placebo) Effect at age 7		-0.016		-0.004
		(0.010)		(0.008)
(Placebo) Effect at age 8		-0.012		-0.001
		(0.012)		(0.009)
Effect at age 9		-0.015		0.015
		(0.012)		(0.010)
Effect at age 10		-0.007		0.009
		(0.013)		(0.011)
Effect at age 11		-0.009		0.002
		(0.013)		(0.011)
Effect at age 12		-0.004		0.014
		(0.013)		(0.011)
N	53,984	53,984	69,752	69,752

Notes: The DD models use retrospective sports club membership information by child age. Columns (1) and (2) use treated and untreated cohorts only from Saxony (6,748 individuals observed at eight different ages) along with age and cohort fixed effects. Columns (3) and (4) use third graders from 2008 to 2010 ("treated cohorts") in Saxony, Brandenburg, and Thuringia (8,719 individuals observed at eight different ages) along with age and state fixed effects. Robust standard errors allowing for clustering at the municipality level are in parentheses (* $p<0.1$, ** $p<0.05$, *** $p<0.01$). The uneven columns show the average effect whereas the even columns show event study estimates by child age.
Source: YOLO survey.

Table B3: Characteristics of Parents of Sports Club Members

Variable	Non-member (1)	Member (2)	Difference (3)	Norm. diff. (4)	N (5)
Highest school degree	0.43	0.53	-0.10***	-0.14	2,800
Middle school degree	0.50	0.44	0.06***	0.09	2,800
Lowest school degree	0.06	0.03	0.03***	0.10	2,800
Age	46.66	47.17	-0.51*	-0.05	2,727
Member of sports club	0.20	0.36	-0.16***	-0.26	2,814
Sport 1	3.43	3.99	-0.56***	-0.19	2,782
Sport 2	3.18	3.46	-0.28***	-0.10	2,761
Sport 3	3.11	3.39	-0.27***	-0.09	2,759
Sport 4	2.99	3.40	-0.41***	-0.14	2,753
Sport 5	3.13	3.40	-0.27***	-0.09	2,735
Sport 6	2.55	2.91	-0.36***	-0.14	2,753
Sport 7	3.99	4.57	-0.58***	-0.17	2,776
Sport 8	3.03	3.56	-0.53***	-0.18	2,753
Sport 9	3.28	3.65	-0.37***	-0.12	2,756
Sport 10	3.14	3.66	-0.51***	-0.17	2,755
Sport 11	3.70	4.43	-0.73***	-0.22	2,763
Exercising index	-0.38	0.29	-0.67***	-0.17	2,655

Notes: The table compares the characteristics of parents according to their children's sports club membership status at age eight. The variables "Sport 1-Sport 11" refer to the responses to 11 statements about exercising. Parents could answer on a Likert Scale from 1 (totally disagree) to 7 (totally agree). The statements are : (1) Exercising is something that I do regularly. (2) Exercising is something that I do automatically. (3) Exercising is something that I do without explicitly reminding myself. (4) Exercising is something that I feel I need if I don't do it. (5) Exercising is something that I do without thinking about it. (6) Exercising is something that would be exhausting for me not to do. (7) Exercising is something that is part of my weekly routine. (8) Exercising is something that would be difficult for me not to do. (9) Exercising is something that I do without the need to think about it. (10) Exercising is something that is typical for me. (11) Exercising is something that I have been doing for a long time.
Source: YOLO survey.

Table B4: Socio-Demographics of Children who Redeemed the Voucher

Variable	Redeemer (1)	Non-redeemer (2)	Difference (3)	Norm. diff. (4)	N (5)
Female	0.57	0.58	-0.01	-0.01	5,025
Has siblings	0.88	0.86	0.02	0.04	4,965
Born in Germany	0.97	0.96	0.01*	0.06	4,993
Parent not born in Germany	0.10	0.15	-0.05***	-0.11	4,995
Newspaper at home	0.64	0.55	0.09***	0.13	4,992
Art at home	0.81	0.73	0.08***	0.13	4,998
Academic track	0.68	0.59	0.09***	0.13	4,997
Sports club at age 4, 5, 6, or 7	0.67	0.51	0.16***	0.23	4,670
Music lessons at age 4, 5, 6, or 7	0.54	0.47	0.07***	0.10	4,670
Number of sports clubs (ZIP code)	16.02	15.64	0.38	0.02	5,027

Notes: The table compares the characteristics of children in treated cohorts in Saxony according to whether they redeemed the voucher or not. Norm. diff. stands for the "normalized difference", which is defined for each variable x as $ND_x = (\overline{x_1} - \overline{x_0}) / \sqrt{(s_{x1}^2 + s_{x0}^2)}$, where $\overline{x_1}$ and $\overline{x_0}$ are the sample means of the two groups and s_{x1}^2 and s_{x0}^2 the corresponding variances. According to Imbens and Wooldridge (2009), a normalized difference of more than 0.25 indicates substantial covariate imbalance.
Source: YOLO survey.

Table B5: Difference-in-Differences: Alternative Methods of Inference

							Wild cluster bootstrap					
			Cluster				Testing under H_0			Testing under H_1		
	rob	conv	municip.	states	twoway	cohort	Radem.	Mammen	Webb	Radem.	Mammen	Webb
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
Panel A: Awareness & take-up												
Program known												
p-value	0.000	0.000	0.000	0.000	0.000	0.014	0.008	0.016	0.005	0.004	0.001	0.000
Voucher received												
p-value	0.000	0.000	0.000	0.000	0.000	0.027	0.019	0.017	0.018	0.000	0.002	0.000
Voucher redeemed												
p-value	0.000	0.000	0.000	0.000	0.000	0.035	0.017	0.028	0.019	0.001	0.002	0.000
Panel B: Physical activity & overweight												
Member of sports club												
p-value	0.631	0.637	0.636	0.224	0.244	0.565	0.372	0.352	0.389	0.346	0.350	0.341
Weekly hours of sport												
p-value	0.991	0.991	0.991	0.977	0.987	0.993	0.990	0.976	0.994	0.990	0.991	0.996
Overweight												
p-value	0.779	0.768	0.795	0.679	0.637	0.774	0.731	0.720	0.747	0.726	0.748	0.745

Notes: The table displays *p*-values for alternative methods of inference using our preferred model specification in column (3) of Table 2. (1) is based on robust standard errors, (2) on conventional standard errors, and (3)-(6) on clustered standard errors, where the level of clustering is the municipality in (3), the state in (4), the cohort and the municipality (two-way clustering) in (5), and the state*cohort group in (6). The *p*-values in (7)-(12) are based on wild cluster bootstrap procedures with state*cohort groups as clusters, where testing is under the null hypothesis in (7)-(9) and under the alternative hypothesis in (10)-(12). (7) and (10) apply Rademacher weights, while (8) and (11) use Mammen weights, and (9) and (12) Webb weights. All wild cluster bootstrap specifications are estimated with the help of the user-written Stata-program BOOTTEST (Roodman et al., 2010).
Source: YOLO survey.

Table B6: Limits of Confidence Intervals

Outcome	$\hat{\beta}$ (1)	95% CI (2)	90% CI (3)
Panel A: Base DD (Table 2, col. (1))			
Member of sports club	0.004	0.035	0.028
	(0.019)		
Weekly hours of sport	-0.069	0.196	0.137
	(0.161)		
Overweight	0.005	-0.021	-0.016
	(0.016)		
Panel B: Main specification (Table 2, col. (3))			
Member of sports club	0.009	0.040	0.033
	(0.019)		
Weekly hours of sport	-0.002	0.260	0.202
	(0.159)		
Overweight	0.004	-0.022	-0.017
	(0.016)		
Panel C: Synthetic control (Table 6, col. (1))			
Member of sports club	-0.008	0.028	0.020
	(0.022)		
Weekly hours of sport	-0.048	0.284	0.211
	(0.202)		
Overweight	-0.017	-0.055	-0.046
	(0.023)		
Panel D: With 2011/2012 cohort (Table 4, col. (3))			
Member of sports club	-0.014	0.011	0.005
	(0.015)		
Weekly hours of sport	-0.148	0.048	0.005
	(0.119)		
Overweight	0.005	-0.011	-0.008
	(0.01)		

Notes: The table displays the 95% (column [2]) and 90% (column [3]) limits of confidence intervals (CI) for the C2SC effect on the outcomes in the first column, based on different specifications as indicated by the panel headers. More specifically, the table shows, based on one-sided tests and the hypothesized sign of the effect, the upper limit of the confidence interval for the outcomes *member of sports club* and *weekly hours of sport* as well as the lower limit for *overweight*.
Source: YOLO survey.

Table B7: Power Calculations

Outcome	Pseudo treatment effect		
	3 units (1)	4 units (2)	5 units (3)
Panel A: Base DD (Table 2, col. (1))			
Member of sports club	0.596	.759	0.902
Weekly hours of sport	0.699	0.863	0.964
Overweight	0.713	0.884	0.974
Panel B: Main specification (Table 2, col. (3))			
Member of sports club	0.596	0.779	0.904
Weekly hours of sport	0.706	0.872	0.968
Overweight	0.668	0.838	0.95
Panel C: Synthetic control (Table 6, col. (1))			
Member of sports club	0.552	0.786	0.91
Weekly hours of sport	0.718	0.902	0.967
Overweight	0.45	0.706	0.916
Panel D: With 2011/12 cohort (Table 4, col. (3))			
Member of sports club	0.722	0.87	0.959
Weekly hours of sport	0.826	0.949	0.992
Overweight	0.916	0.977	1

Notes: The table shows the statistical power of different DD model specifications as indicated by the panel headers. The simulations use data from the YOLO survey, artificially induce pseudo treatment effects as indicated by the column headers and 1,000 replications. First, we randomly assign 30 municipalities—the same number that is actually treated—to the treatment group and the other municipalities to the control group. Then we artificially introduce treatment effects of different magnitudes in indicated by the column headers. Next, we run each DD model 1,000 times. Based on the hypothesized sign of the effect sizes, we add increases of 3, 4, and 5 percentage points as pseudo treatment effects for *member of sports club*; for *weekly hours of sport*, we add increases of 0.3, 0.4. and 0.5 hours; and for *overweight*, we add *de*creases of 3, 4, and 5 percentage points. Each cell indicates the share of the 1,000 replications for which the models reject the null hypothesis of statistically significant effects at the 10 percent significance level in one-sided tests. For example, the simulation for the second row of Panel A in column (2) indicates that our base DD model in Table 2 can correctly identify an increase of 0.4 weekly hours of sport at the 90 percent certainty level in 86 percent of all 1,000 replications.
Source: Simulations based on YOLO survey.

Appendix C: School Examination Data

Table C1: Descriptive Statistics—School Health Examination Data

Variable	Mean	Std. Dev.	Min.	Max.	N
	(1)	(2)	(3)	(4)	(5)
Age in months	144.993	7.070	120	183	6,794
Female	0.502	0.5	0	1	6,794
Height in cm	154.236	8.158	127	186	6,794
Weight in kg	46.108	11.471	22	140	6,794
BMI	19.2026	3.636	10.54	42.31	6,794
Obese	0.0596	0.2377	0	1	6,794
Overweight	0.1099	0.3129	0	1	6,794
Underweight	0.1097	0.3125	0	1	6,794
Hypertension	0.078	0.268	0	1	6,794
Motor skill disorder	0.004	0.065	0	1	6,794
Emotional Disorder	0.036	0.186	0	1	6,794
Weak posture	0.115	0.319	0	1	6,794

The table shows descriptive statistics for our administrative School Health Examination Data from one county in Saxony. "Std. Dev." stands for standard deviation and "N" indicates the number of unique students in our data. *Source*: School Health Examination Data from Public Health Service (*Öffentlicher Gesundheitsdienst*).

Table C2: Regression of Objective Health Outcomes on Treated and Control Cohorts

	Obese (1)	Overweight (2)	Motor skill disorder (3)	Emotional disorder (4)
Treated	-0.014	0.007	-0.002	0.000
	(0.014)	(0.005)	(0.002)	(0.014)
Age in months	0.002***	0.000	0.000	0.002***
	(0.000)	(0.001)	(0.000)	(0.000)
Female	0.007	-0.019**	-0.001	-0.025***
	(0.008)	(0.008)	(0.001)	(0.005)
Month of examination	X	X	X	X
Observations	6,794	6,794	6,794	6,794
R-squared	0.003	0.002	0.001	0.007

Notes: The table shows four separate regressions where the column headers indicate the objective health data used as outcome variable. *Treated* is a dummy indicating whether the cohort was treated whereas *age in months* and *female* are control variables, see main text for details. Robust standard errors in parentheses are clustered at the examination date level. *** $p<0.01$, ** $p<0.05$, * $p<0.1$.
Source: School Health Examination Data from Public Health Service (*Öffentlicher Gesundheitsdienst*).