

The Long-Run Effects of Sports Club Vouchers for Primary School Children[†]

By JAN MARCUS, THOMAS SIEDLER, AND NICOLAS R. ZIEBARTH*

Childhood obesity is one of the most serious public health challenges of the twenty-first century. While small-scale experiments change behaviors among adults in the short run, we know little about the effectiveness of large-scale policies or the longer-run impacts. To nudge primary school children into a long-term habit of exercising, the German state of Saxony distributed sports club membership vouchers among all 33,000 third graders in 2009. In 2018, we carried out a register-based survey to evaluate the policy. 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. Apparently, membership vouchers for children are not a strong enough policy tool to overcome barriers to exercise regularly. (JEL H75, I12, I18, I21, I28, J13, Z21)

According to the World Health Organization (2020), childhood obesity is “one of the most serious public health challenges of the 21st century.” To fight childhood obesity, 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, Dietz, and Collins 2010; World Health Organization 2017). However, we know very little about which policies actually

*Marcus: School of Business & Economics, Freie Universität Berlin, Department of Economics, Universität Hamburg, IZA, and DIW Berlin (email: jan.marcus@fu-berlin.de); Siedler: Universität Potsdam and IZA, Department of Economics (email: thomas.siedler@uni-potsdam.de); Ziebarth: Department of Economics, Brooks School of Public Policy, Cornell University, IZA, NBER, and ZEW Mannheim (email: nrz2@cornell.edu). C. Kirabo Jackson was coeditor for this article. We would like to thank Sascha O. Becker, John Cawley, Jan Feld, Paola Giuliano, 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 Sächsisches Staatsministerium für Kultus, Mr. Ender from the Kultusministerium Sachsen, Ms. Riedel from the Sozialministerium Sachsen, and Ms. Frenschkowski from the Sächsisches Staatsministerium für 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.

[†]Go to <https://doi.org/10.1257/pol.20200431> to visit the article page for additional materials and author disclosure statements or to comment in the online discussion forum.

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, each 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 was 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 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 19,000 youth completed the survey. We use registry data to compare characteristics of survey participants and nonparticipants and show that YOLO 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 dataset, we exploit that state-employed physicians are legally required to physically examine

¹Cawley, Meyerhoefer, and Newhouse (2007) and Cawley, Frisvold, and Meyerhoefer (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, Frisvold, and Meyerhoefer 2013) and an impact on actual PE instruction time in high school, but no significant effect on high school students' body weight (Cawley, Meyerhoefer, and Newhouse 2007). Otherwise, the rich economics obesity literature focuses on adult obesity and studies general driving forces such as technological change (Cutler, Glaeser, and Shapiro 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, Frisvold, and Tefft 2010), increases in portion sizes (Jeitschko and Pecchenino 2006), increases in gluttony (Griffith, Lluberas, and Lührmann 2016), declining gas prices (Courtemanche 2011), increasing cigarette taxes (Courtemanche 2009), changes in food prices (Dubois, Griffith, and Nevo 2014; Grossman, Tekin, and Wada 2014; Strulik 2014; Courtemanche, Heutel, and McAlvanah 2015; Dragone and Ziebarth 2017), and cash transfers (Akee et al. 2013).

all children in school. We 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. 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, James, and Nolen 2016; Loewenstein, Price, and Volpp 2016). Second, as the vouchers 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, Stehr, and Sydnor 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, Feiler, and Wicker 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 purpose other 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, Stehr, and Sydnor 2015; Carrera et al. 2018, 2020). Exceptions are Reichert (2015) and Augurzyk 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).

The next section explains the extramural sports club setting in Germany and describes the voucher policy in detail. Section II discusses the data and key variables, and Section III explains the empirical models. Section IV presents our findings based on YOLO, and Section V presents robustness checks. Section VI discusses additional evidence based on official health examination data, and Section VII concludes.

I. Nonprofit Sports Clubs and the Voucher Initiative

A. 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, nonprofit 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,

²In 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.

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

⁴According 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 nonprofit 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.

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 that young people in the US pursue when they join school-based sports teams. As in the United States, 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 are an integral part of the experience that young people gain when they participate in extramural sport.

B. *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. We provide an example of the voucher, in the accompanying working paper (Marcus, Siedler, and Ziebarth 2021). 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 email 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, policy-makers 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).

II. 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 (Marcus, Siedler, and Ziebarth 2022). This main dataset, the Youth Leisure Online Survey (YOLO), uses a two-stage sampling strategy, which we describe in detail in the online Appendix and Figure A1. 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

⁵In 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 V).

⁶Note 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.

⁷While 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.

proportional to population size).⁸ Online Appendix Figure A2 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. Online Appendix Table A1 lists the relationship between birth cohorts and school cohorts. We aimed to survey two prevoucher 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

Next, we mailed one official invitation letter to each of the 155,527 adolescents sampled in the second stage. Online Appendix Figure A3 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 online Appendix Figure A4). 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 VI, in addition to YOLO, we use official student health examination data carried out by state-employed physicians. We use this second dataset from one county in Saxony to examine objectively measured body weights, heights, as well as motor skill and emotional disorders.

A. 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

⁸We 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. Ninety-four 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.

⁹Note 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 IB). Second, this is the first cohort that did not receive vouchers and, therefore, disappointment effects could arise. However, in Section V we show that our findings are robust to including this post-treatment cohort.

¹⁰The lottery prize included two iPads, worth €500 each, and ten Amazon vouchers, worth €20 each.

¹¹We invited one parent (mother or father) to participate in the survey.

(*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 online Appendix Figure A3). It only stated that the survey would be about leisure time behavior among young people. Further, so as not to prime 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.

B. 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, we work with two pretreatment 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.

C. Descriptives

Table 1 shows the summary statistics of the YOLO sample. The average age of the young people surveyed was 17.5 years. Fifty-seven percent of all respondents were female. About one-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 socioeconomic 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.

Thirty-eight 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 10 percent recall having received the voucher ("Voucher received") and 6 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." Online Appendix Figure B1 plots the sports disciplines for which the vouchers were redeemed. Not surprisingly, by far the most popular discipline was

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

TABLE 1—SUMMARY STATISTICS

Variable	Mean (1)	SD (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. Observations = 13,334.

Source: YOLO survey

soccer (20 percent), followed by martial arts (12 percent), swimming (12 percent), handball (7 percent), athletics (6 percent), and gymnastics (5 percent).

Online Appendix Table A2 lists the normalized difference for key variables between treatment and control groups. According to Imbens and Wooldridge (2009), a normalized difference of 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) versus 17.4 (control); about 57 percent of respondents were female in both groups, and 20.5 percent versus 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.

D. 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 (German Registry Offices 2017) and the German Socio-Economic Panel Study, SOEP (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

¹³One of the very few covariates with a major imbalance is the share of youth living in a city (76 percent in the treatment group versus 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.

participants were not aware that we had the registry information. Online Appendix Table A3 shows that 99 percent of all YOLO participants correctly reported their gender, nationality, and year of birth.¹⁴

Second, based on the registry data, online Appendix Table A4 compares the characteristics of YOLO participants and nonparticipants. While YOLO has a slight overrepresentation of women, Germans, and younger individuals, all normalized differences are below 0.25, suggesting that participants and nonparticipants do not differ strongly with respect to these characteristics. Additionally, in Section V, 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 dataset 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 we designed the wording of several YOLO questions to be 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 1997 and July 2000 and who live in Saxony, Brandenburg, or Thuringia.¹⁵ Online Appendix Table A5 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.

III. Empirical Approach

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

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

where Y_{ics} denotes the outcome of individual i in cohort c in state s . The dummy variable $Saxony_s$ is 1 if a respondent was a third grader in Saxony, and 0 if he or she was a third grader in Thuringia or Brandenburg. Recall that about 33,000 third graders of the treatment cohorts were treated and received the voucher. $Post_c$ is a dummy variable, which is 1 if a respondent was a third grader in the school year 2008/09, 2009/10 or 2010/11, and 0 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

¹⁴Four 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 online Appendix Table A3. In the robustness section, we show that the results are robust to excluding individuals with nonmatching registry information and individuals who spent little time answering the survey.

¹⁵We 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.

two dummies (*Saxony · Post*). It is 1 if a respondent was eligible to receive and redeem sports club vouchers, and 0 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 two-way 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:

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

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.

IV. Results

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

A. 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. Panels A–C of Figure 1 illustrate whether treated cohorts in Saxony have a higher awareness of the program (panel A), were more likely to have received the voucher (panel B), and were more likely to have redeemed the voucher (panel C).

As seen, we observe very flat and nontrending lines over the entire time period for respondents in the control states for all three outcomes (dashed lines).¹⁷ In contrast,

¹⁶In Section V, we present and discuss several alternative methods of inference.

¹⁷In panel A of Figure 1, some individuals in Saxony's pretreatment cohorts report to have known the program, and in panels B and C a small fraction of individuals in Saxony's pretreatment 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

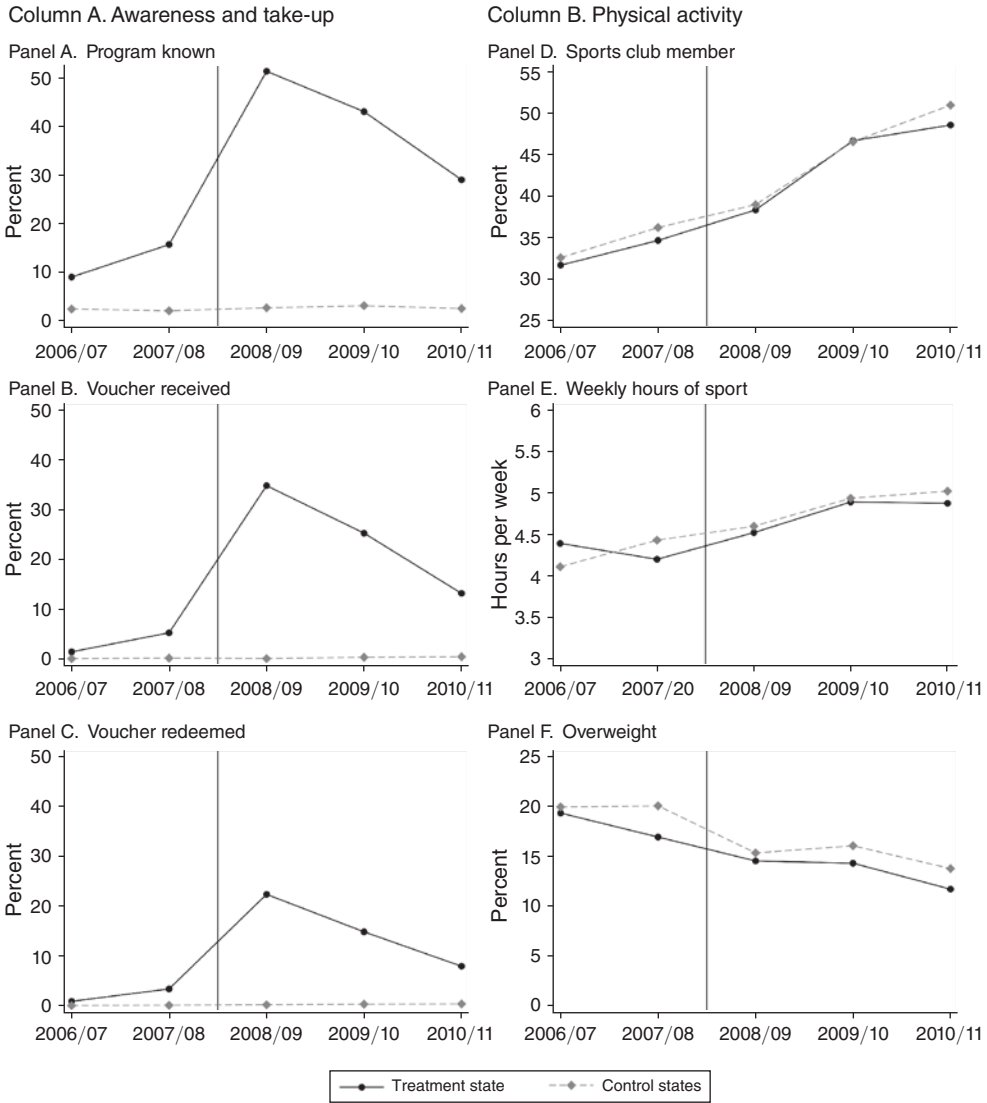


FIGURE 1. DEVELOPMENT OF OUTCOME VARIABLES IN TREATMENT AND CONTROL STATES ACROSS COHORTS

Notes: The figure displays 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

respondents who were third graders in Saxony at the time of the policy show a substantially larger program 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

out the voucher to the wrong class), or that individuals remember incorrectly whether they received the voucher (recall bias).

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 panels B and C of Figure 1: there are substantial spikes in the first year of the voucher initiative, and subsequent linear decreases in years 2 and 3. 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 IB for details).

Panels D, E, and F of Figure 1 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 1 for respondents who attended third grade in Saxony), a binary post-reform indicator (which is 1 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.

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

TABLE 2—EVALUATION OF SPORTS CLUB VOUCHER PROGRAM: MAIN DD RESULTS

	Base DD (1)	Two-way fixed effects (2)	+ municip. fixed effects (3)
<i>Panel A. Awareness and take-up</i>			
Program known			
Voucher	0.272 (0.014)	0.272 (0.014)	0.276 (0.014)
Voucher received			
Voucher	0.200 (0.011)	0.201 (0.011)	0.202 (0.011)
Voucher redeemed			
Voucher	0.122 (0.006)	0.122 (0.006)	0.122 (0.006)
<i>Panel B. Physical activity and overweight</i>			
Member of sports club			
Voucher	0.004 (0.019)	0.003 (0.019)	0.009 (0.019)
Weekly hours of sport			
Voucher	-0.069 (0.161)	-0.082 (0.159)	-0.002 (0.159)
Overweight			
Voucher	0.005 (0.016)	0.006 (0.016)	0.004 (0.016)
Observations	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. Each cell 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 1 for respondents who attended third grade in Saxony in school years 2008/2009, 2009/2010, or 2010/2011. 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

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 V). (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.

¹⁸For the outcome voucher received we would expect a point estimate close to 1 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).

B. 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 subgroups. Technically, we run a version of equation (2) with municipality fixed effects, to which we add interactions of all variables with the stratifying variable of interest:

$$(3) \quad Y_{igcs} = \beta_1 \cdot \text{Voucher}_{cs} + \beta_2 \cdot \text{Voucher}_{cs} \cdot \text{Group}_{igcs} + \gamma_{cg} + \kappa_{sg} \\ + \lambda \cdot \text{Group}_{igcs} + \varepsilon_{igcs},$$

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 1 if an individual belongs to a specific group (e.g., females) and 0 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, online Appendix Table B1 shows the DD regression coefficients for all subgroups.¹⁹

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, online Appendix Table B1 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. Online Appendix Table B1 shows no single significant treatment effect for these long-term outcomes for any of the subgroups.

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 non-parametric evidence from Figure 1: the highly significant effects on awareness,

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

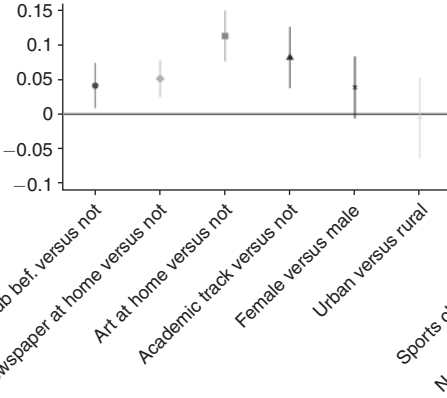
²⁰For this purpose, we add cohort-specific treatment indicators to equation (2) with municipality fixed effects:

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

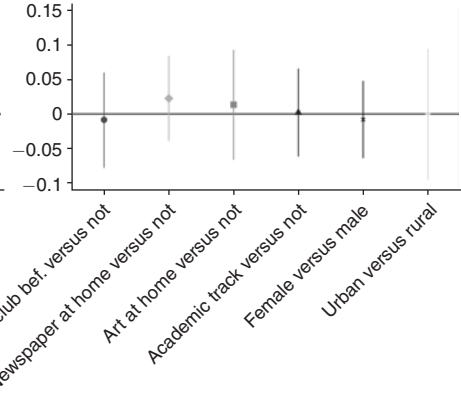
Panel A. Awareness and take-up

Panel B. Physical activity

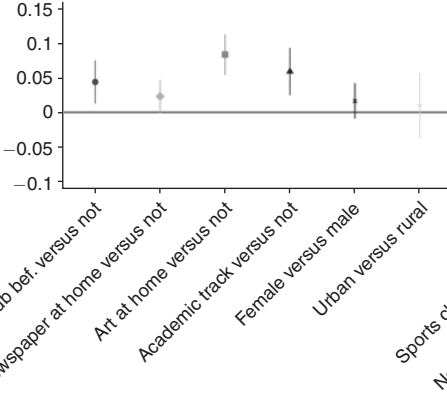
(a) Program known



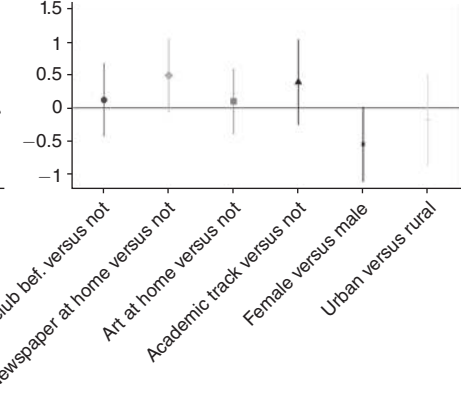
(d) Sports club member



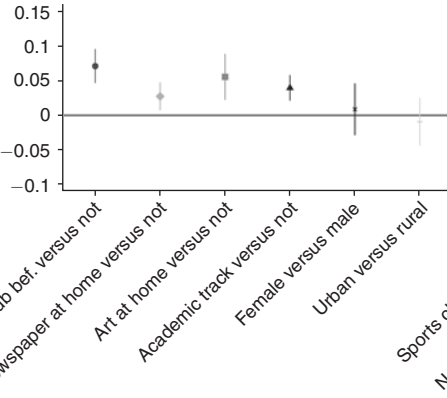
(b) Voucher received



(e) Weekly hours of sport



(c) Voucher redeemed



(f) Overweight

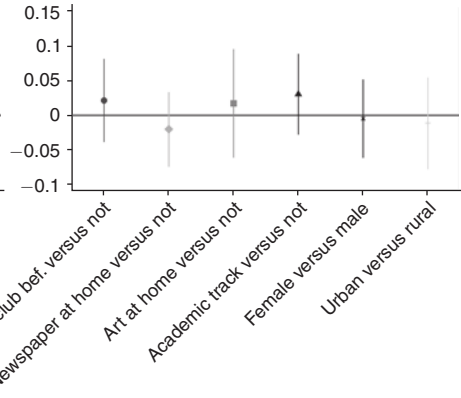


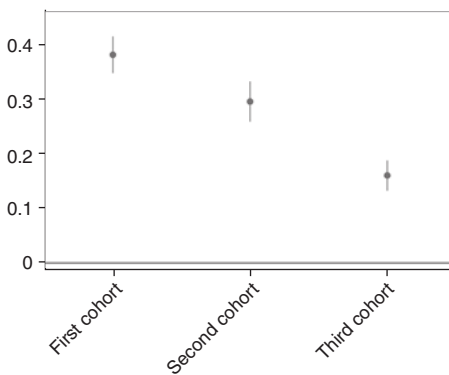
FIGURE 2. EFFECT HETEROGENEITY ACROSS SOCIO-DEMOGRAPHICS

Note: The figure displays 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

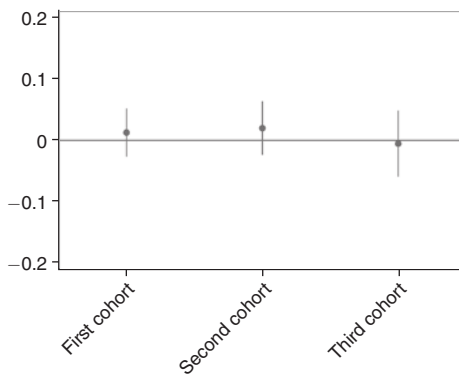
Panel A. Awareness and take-up

(a) Program known

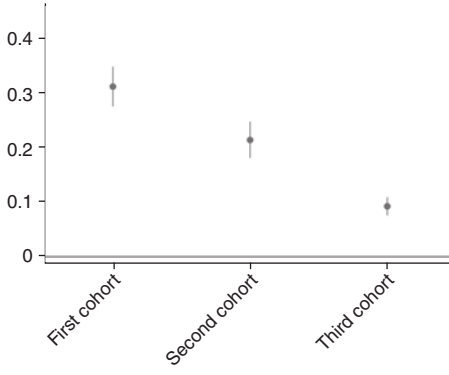


Panel B. Physical activity

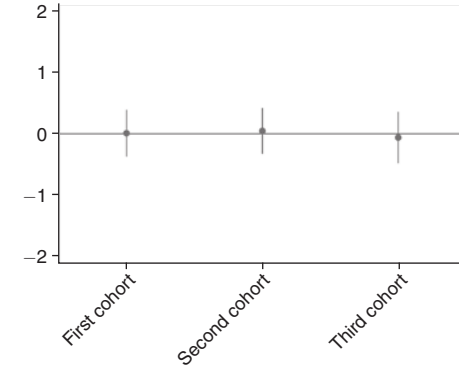
(d) Sports club member



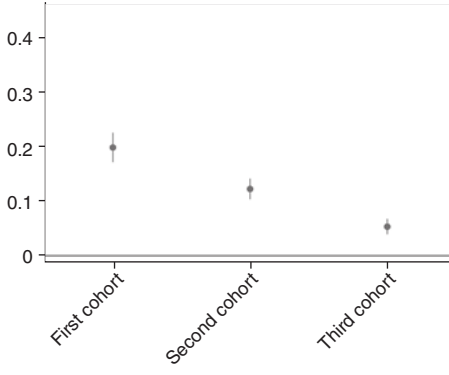
(b) Voucher received



(e) Weekly hours of sport



(c) Voucher redeemed



(f) Overweight

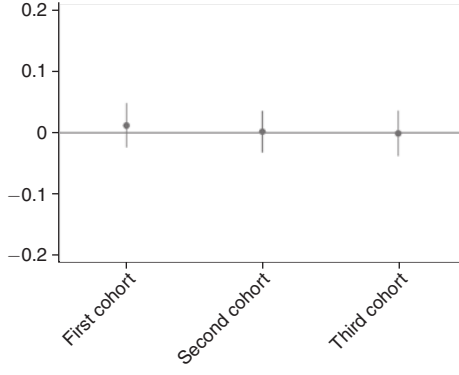


FIGURE 3. EFFECT HETEROGENEITY ACROSS COHORTS

Note: The panels 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

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.

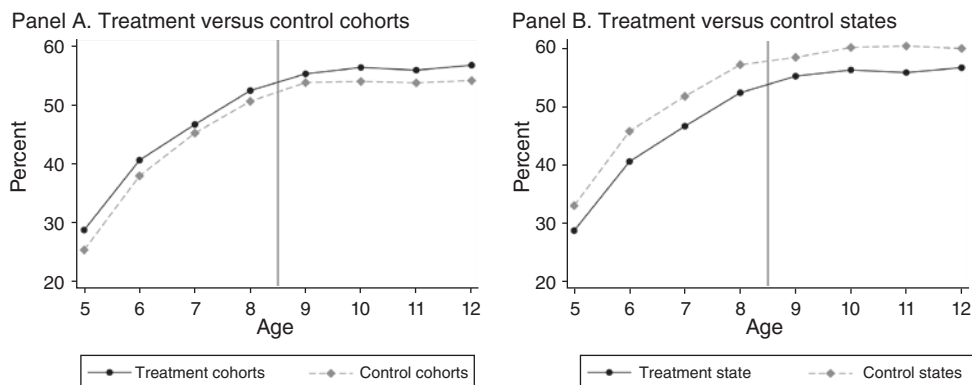


FIGURE 4. SPORTS CLUB MEMBERSHIP BY AGE

Notes: The panels display retrospectively reported sports club membership rates by child age. Panel 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). Panel 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

C. 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.

Panel A of Figure 4 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 nontreated 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 panel B of Figure 4, 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 online Appendix Table B2. 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 panel A of Figure 4 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 mirror the visual evidence in panel B of Figure 4 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

Online Appendix 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 nonsignificant 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.

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 nontreated cohorts, providing support for this argument.

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

$$Y_{ica} = \beta \cdot \text{Voucher}_{ca} + \gamma_c + \mu_a + \varepsilon_{ica}$$

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 1 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:

$$Y_{isa} = \beta \cdot \text{Voucher}_{sa} + \kappa_s + \mu_a + \varepsilon_{isa}$$

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 panel A of Figure 4 (and columns 1 and 2 in online Appendix 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 panel B (and columns 3 and 4 in Table B2) is that the recall bias among those who were third graders in 2008/09 to 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.

TABLE 3—SPORTS CLUB MEMBERSHIP ACROSS CHILD AGES

	Children		Parents' response (3)
	All (reference) (1)	Parent sample (2)	
Member of sports club at age 6			
Voucher	-0.019 (0.016)	0.032 (0.046)	-0.017 (0.051)
Member of sports club at age 7			
Voucher	-0.025 (0.019)	0.014 (0.044)	0.016 (0.045)
Member of sports club at age 8			
Voucher	-0.023 (0.021)	0.004 (0.060)	0.032 (0.053)
Member of sports club at age 9			
Voucher	-0.007 (0.018)	0.077 (0.056)	0.011 (0.049)
Member of sports club at age 10			
Voucher	0.014 (0.015)	0.061 (0.054)	-0.010 (0.047)
Member of sports club at age 11			
Voucher	0.003 (0.018)	0.029 (0.058)	-0.023 (0.056)
Member of sports club at age 12			
Voucher	0.008 (0.016)	-0.023 (0.052)	-0.025 (0.057)
Observations	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 columns 2 and 3 directly comparable. Robust standard errors allowing for clustering at the municipality level are in parentheses. Each column in each panel represents one DD estimate. Models are based on equation (2) with municipality fixed effects.

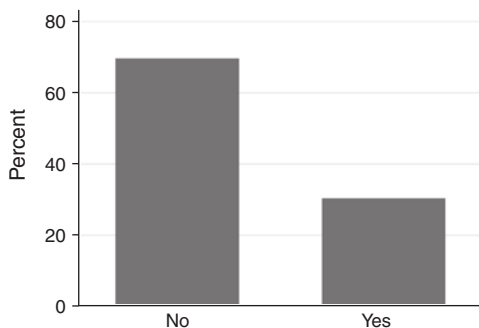
Source: YOLO survey

Second, to affect the consistency of our estimates, recall bias would have to take a nontrivial 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 does not elicit 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 panel A of Figure 5).

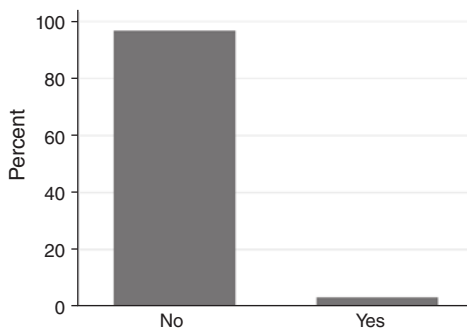
D. Suggestive Evidence on Mechanisms

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

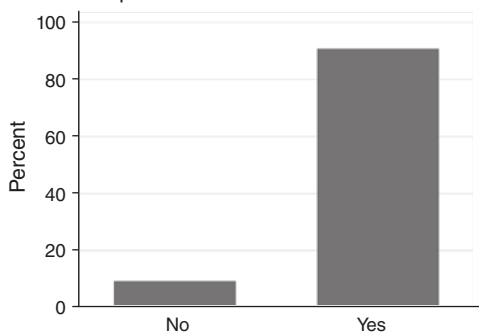
Panel A. Tried new sport discipline(s)



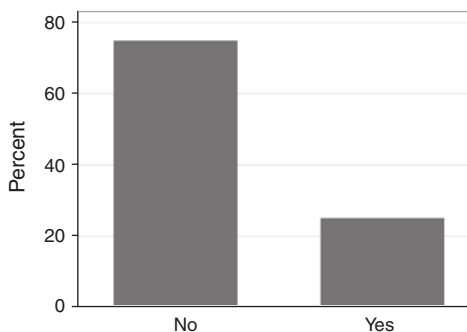
Panel C. Could not afford membership without voucher



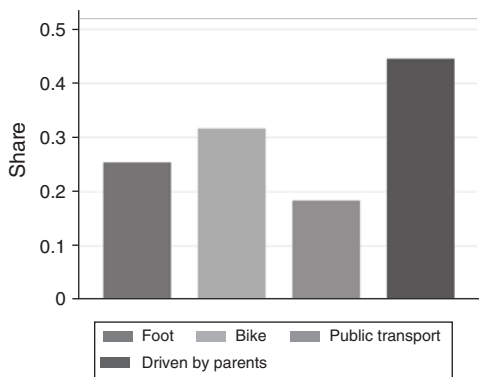
Panel B. Could redeem voucher for desired discipline



Panel D. Parents happy to save money because of voucher



Panel E. Transport as a supply-side barrier



Panel F. Mode of transportation (urban versus rural)

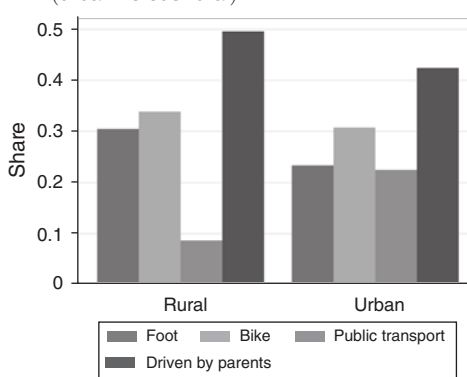


FIGURE 5. SUGGESTIVE EVIDENCE ON MECHANISMS

Notes: Panels 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. Panels 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

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. Panel B of Figure 5 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, panel A of Figure 6 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 6 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 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.

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 panel F of Figure 5. As seen, children in rural counties were much more likely to get a ride from their parents (50 versus 42 percent). However, they were also more likely to walk (30 versus 23 percent) or bike (34 versus 30 percent) to the sports club. Finally, not surprisingly, they were significantly less likely to take public transport (9 versus 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.

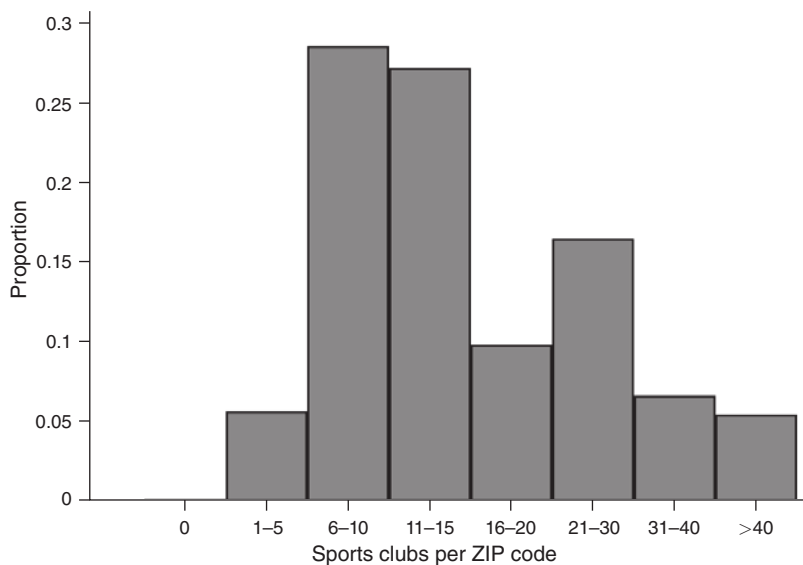
In sum, this suggests that supply-side constraints were unlikely to be a major barrier to take-up.

²⁴These 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>).

²⁵Technically, 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).

²⁶Unlike in the United States, “free range parenting” is the social norm in Germany.

Panel A. Distribution of the number of sports clubs



Panel B. Distribution of the number of sports club disciplines

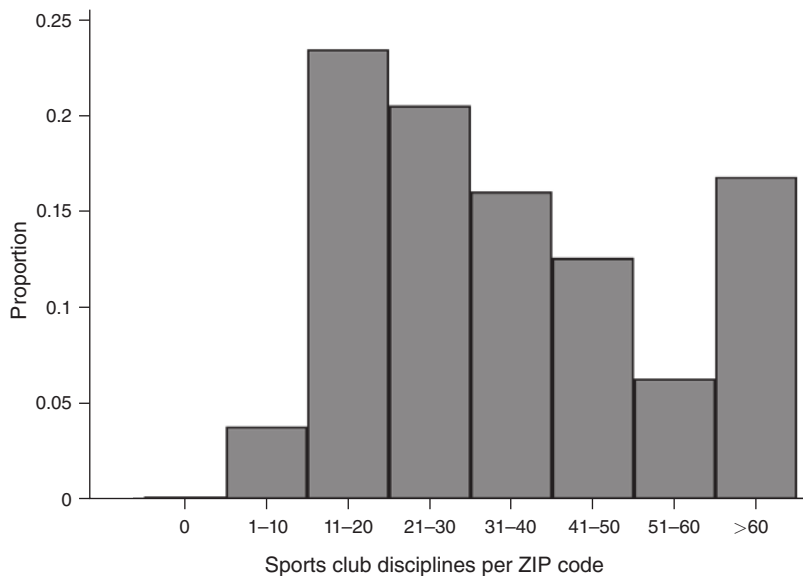


FIGURE 6. SUPPLY-SIDE RESTRICTIONS? NUMBER OF SPORTS CLUBS PER ZIP CODE

Notes: The panels 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

Demand-Side Arguments.—First, it could be that the program targeted the wrong age group. However, as almost one-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).

Second, stigma can be a reason for incomplete take-up of social benefits (Friedrichsen, König, and Schmacker 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 IB). 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. Online Appendix Table B3 compares the characteristics of parents of children who are versus 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 a similar exercise but focus on the treated cohorts in Saxony and differentiate between children who redeemed versus those who did not redeem the voucher. This time we use responses from the youth questionnaire. The findings in online Appendix Table B4 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

²⁷We 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.

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

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 nonmonetary 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, panels C and D of Figure 5 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, one-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 nonmonetary 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.

V. Sensitivity Analyses

A. *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). Table 5 presents the results of a specification that only considers individuals with matching registry information (regarding gender, nationality, and age) and who spent at least ten minutes on the survey (column 1). In column 2, to consider potential spillover effects between siblings, we exclude respondents from the pretreatment cohorts in all three states who reported that their siblings also received an invitation to participate in YOLO. In column 3, we exclude youth with an older sibling.

TABLE 4—ROBUSTNESS: TEMPORAL AND REGIONAL RESTRICTIONS

	Temporal				Regional		
	2007–2010 (1)	2000–2010 (2)	2006–2011 (3)	2006–2009 (4)	Without BB (5)	Without TN (6)	All (7)
<i>Panel A. Awareness and take-up</i>							
Program known	0.245	0.279	0.265	0.337	0.271	0.279	0.263
Voucher	(0.011)	(0.014)	(0.011)	(0.017)	(0.017)	(0.014)	(0.016)
Voucher received	0.186	0.203	0.196	0.262	0.202	0.202	0.196
Voucher	(0.010)	(0.011)	(0.010)	(0.017)	(0.012)	(0.011)	(0.011)
Voucher redeemed	0.111	0.123	0.118	0.159	0.122	0.122	0.117
Voucher	(0.006)	(0.006)	(0.006)	(0.010)	(0.006)	(0.006)	(0.006)
<i>Panel B. Physical activity and overweight</i>							
Member of sports club	0.011	0.006	−0.014	0.016	0.017	0.004	0.010
Voucher	(0.019)	(0.019)	(0.015)	(0.017)	(0.024)	(0.022)	(0.020)
Weekly hour of sports	0.177	−0.054	−0.148	0.052	0.086	−0.034	−0.098
Voucher	(0.179)	(0.164)	(0.119)	(0.162)	(0.185)	(0.189)	(0.154)
Overweight	0.014	0.010	0.005	0.006	0.018	−0.004	0.000
Voucher	(0.022)	(0.015)	(0.010)	(0.016)	(0.022)	(0.018)	(0.015)
Observations	11,686	13,506	16,082	10,044	9,572	10,973	14,720

Notes: The table imposes different temporal and regional restrictions to test the robustness of the C2SC effects for all main outcomes. Robust standard errors allowing for clustering at the municipality level are in parentheses. 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. The column headers indicate the type of robustness check, see main text for details.

Source: YOLO survey

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 4) and use the current state of residence (column 5).²⁹ Column 6 controls for socio-demographics,³⁰ while column 7 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 8 of Table 5). All point estimates are small in size, also the first three on awareness and utilization, and not statistically different from zero.

²⁹Note that geographic mobility in Germany is much lower than in the United States. 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 1 percent of all respondents move (Jürges, Reinhold, and Salm 2011; Goebel et al. 2015).

³⁰These 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 online Appendix Table A2).

³¹More 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.

TABLE 5—FURTHER ROBUSTNESS CHECKS

	Individual			Treatment		Controls		Placebo
	Serious (1)	Sibling (2)	Without older sibling (3)	First grade (4)	Current (5)	+ X_{ces} (6)	Weighted (7)	BB versus TN (8)
<i>Panel A. Awareness and take-up</i>								
Program known	0.282 (0.015)	0.279 (0.017)	0.266 (0.021)	0.264 (0.015)	0.279 (0.015)	0.274 (0.015)	0.268 (0.017)	−0.007 (0.009)
Voucher received	0.205 (0.013)	0.195 (0.011)	0.199 (0.014)	0.196 (0.014)	0.197 (0.012)	0.201 (0.012)	0.200 (0.012)	0.000 (0.004)
Voucher redeemed	0.124 (0.006)	0.115 (0.006)	0.118 (0.007)	0.116 (0.008)	0.119 (0.006)	0.120 (0.007)	0.123 (0.008)	0.001 (0.003)
<i>Panel B. Physical activity and overweight</i>								
Member of sports club	0.014 (0.021)	0.017 (0.020)	0.004 (0.034)	0.010 (0.017)	−0.001 (0.017)	0.017 (0.020)	0.005 (0.020)	0.018 (0.028)
Weekly hour of sports	−0.027 (0.174)	0.029 (0.169)	−0.151 (0.205)	0.030 (0.130)	0.096 (0.139)	0.125 (0.154)	0.026 (0.188)	0.149 (0.204)
Overweight	0.005 (0.019)	−0.002 (0.016)	0.031 (0.029)	0.006 (0.014)	0.005 (0.016)	0.001 (0.018)	−0.002 (0.016)	0.022 (0.024)
Observations	10,836	12,485	5,891	13,421	13,859	12,106	11,993	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. 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 4 and 5). The column headers indicate the type of robustness check, see main text for details. Column 1: “Serious” conditions on respondents who spent at least ten minutes on the survey and whose registry and self-reported information match exactly. Column 2 excludes respondents from the pre-treatment cohorts whose siblings also received a YOLO invitation. Column 3 excludes youth with an older sibling.

Source: YOLO survey

Second, panel A of Figure 7 plots point estimates from our preferred model along with 95 percent confidence bands. However, here we use outcomes that either cannot plausibly be affected 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 ages 4 and 11) 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.

Panel B of Figure 7 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.

In addition to weighting the regressions with the respondents’ probability of participating in YOLO (column 7 of Table 5), online Appendix Table A6 shows that individuals in treated cohorts in Saxony are just as likely as everybody else to participate in YOLO. Using only registry data, we rerun our main DD model in equation

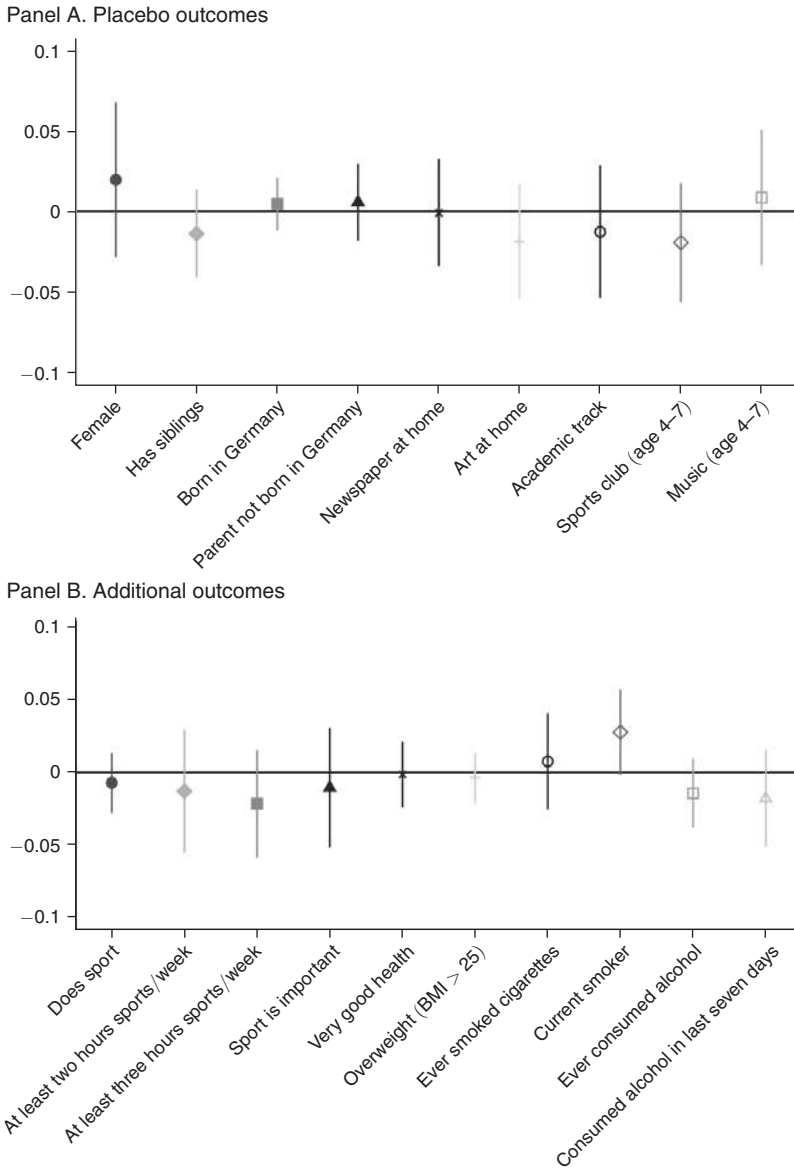


FIGURE 7. FURTHER OUTCOMES

Note: The panels display C2SC treatment effects on placebo outcomes (panel A) and additional outcomes (panel B), using equation (2) with municipality fixed effects.

Source: YOLO survey

(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. Online Appendix Table B5 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.

B. Parents' Information

We also invited parents to fill out a similar survey, which we electronically linked to their children's responses (see Section II). Now, we leverage this information to check whether the results are robust to using parents' responses. Table 6 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 6), while the utilization effect is slightly larger when using the parents' responses.

C. Synthetic Control Group

Although the pretreatment trends in Figure 1 are fairly similar, as another test, we follow Abadie, Diamond, and Hainmueller (2010) and construct a synthetic control group with the same pretreatment outcome trends. More specifically, we reweight the municipalities in the control states such that the two pretreatment 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 7 nor the visual inspections in online Appendix Figure B2 provide evidence that the C2SC initiative increased membership rates and the amount of physical exercise, or reduced overweight rates.

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 online Appendix 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.

D. Statistical Power

As a final exercise, we conduct two types of power analyses. The results are in online Appendix 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 3) and 95 percent (column 2) confidence intervals. For example, the model in the first row of panel A in column 3 indicates that increases in

³²Note that we did not ask parents about their children's current sports activities or weight.

³³The notes for Table 7 provide further details on the construction of the synthetic control group.

TABLE 6—USING PARENTS' RESPONSES FOR PROGRAM-RELATED OUTCOMES

	Children		Parents' response (3)
	All (reference) (1)	Parent sample (2)	
Program known			
Voucher	0.276 (0.014)	0.322 (0.017)	0.271 (0.024)
Voucher received			
Voucher	0.202 (0.011)	0.233 (0.017)	0.275 (0.025)
Voucher redeemed			
Voucher	0.122 (0.006)	0.144 (0.011)	0.167 (0.021)
Observations	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 columns 2 and 3 directly comparable. Robust standard errors allowing for clustering at the municipality level are in parentheses. Models are based on equation (2) with municipality fixed effects.

Source: YOLO survey

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 reinclude 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, online Appendix 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 1,000 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

TABLE 7—SYNTHETIC CONTROL GROUP RESULTS

	Synthetic control	
	Version 1 (1)	Version 2 (2)
<i>Panel A. Awareness and take-up</i>		
Program known		
Voucher	0.285 (0.017)	0.280 (0.017)
Voucher received		
Voucher	0.207 (0.014)	0.203 (0.014)
Voucher redeemed		
Voucher	0.127 (0.007)	0.123 (0.008)
<i>Panel B. Physical activity and overweight</i>		
Member of sports club		
Voucher	-0.008 (0.022)	-0.016 (0.021)
Weekly hours of sport		
Voucher	-0.048 (0.202)	-0.206 (0.290)
Overweight		
Voucher	-0.017 (0.023)	-0.012 (0.021)
Observations _{cohort-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 predetermined 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. Observations refers to the number of cohort-municipality observations. Robust standard errors are in parentheses and clustered at the municipality level.

Source: YOLO survey

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 2021). 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.

VI. 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, measures the height and weight of all children, checks immunizations, and tests 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 (German Public Health Service 2018). 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 versus 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) cutoffs to determine overweightness and obesity for adults do not apply to children. Guidelines determine overweightness and obesity in percentiles 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) and Kromeyer-Hauschild, Moss, and Wabitsch (2015). We use the diagnoses recorded by the state-employed physicians. The guidelines define children as overweight if they fall between the 90th and 97th percentile of the age-specific BMI distribution, as obese if they exceed the 97th percentile, and as underweight if they fall below the 10th percentile of the age-specific BMI distribution. The major advantage of these data is the objective nature of the health and motor skills measures.

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 one missing for BMI per year for a cohort of about 2,500 children.

Online Appendix Table C1 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 6 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.

Panels A to D of Figure 8 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 predetermined control variables in online Appendix Table C2. Overall, the supplemental school health examination data confirm our findings based on the YOLO survey.

VII. 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 Commission 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, Frisvold, and Meyerhoefer 2013), improved educational performance (Stevenson 2010; Fricke, Lechner, and Steinmayr 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 that people start being physically active, the stronger the long-term effects (Hallal et al. 2006; Lally and Gardner 2013; Belot, James, and Nolen 2016; Loewenstein, Price, and Volpp 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

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

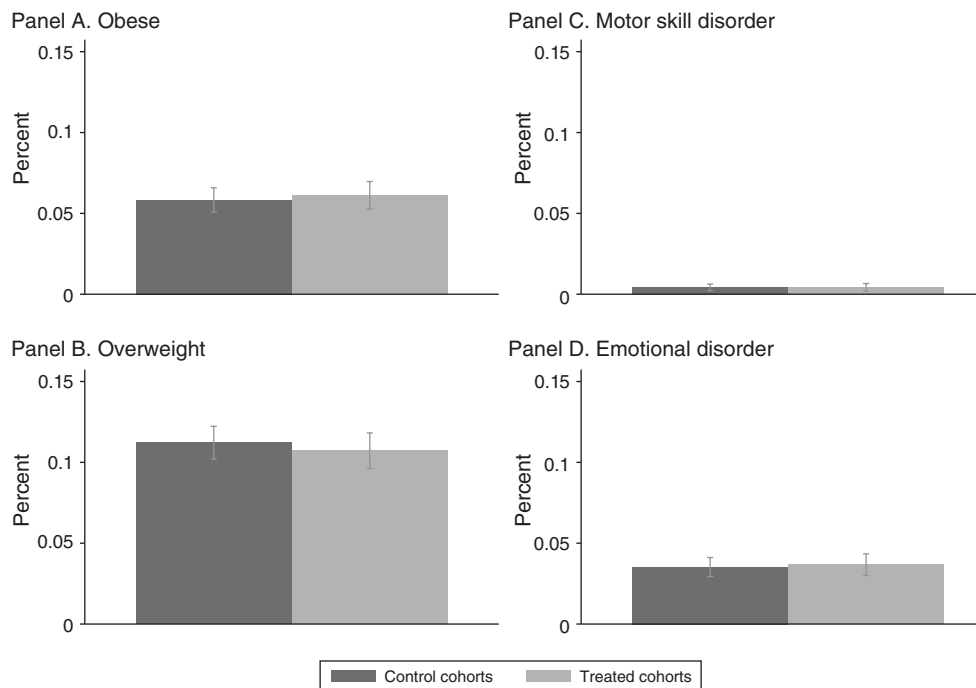


FIGURE 8. OBJECTIVE HEALTH OUTCOMES: TREATMENT VERSUS CONTROL COHORTS

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*)

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 still recall the initiative a decade later. 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 ineffectiveness 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 nonmonetary 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 backgrounds 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 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, Alberto, Susan Athey, Guido W. Imbens, and Jeffrey Wooldridge.** 2017. "When Should You Adjust Standard Errors for Clustering?" NBER Working Paper 24003.
- Abadie, Alberto, Alexis Diamond, and Jens 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.
- Akee, Randall, Emilia Simeonova, William Copeland, Adrian Angold, and E. Jane Costello.** 2013. "Young Adult Obesity and Household Income: Effects of Unconditional Cash Transfers." *American Economic Journal: Applied Economics* 5 (2): 1–28.
- Anderson, Michael L., and David A. Matsa.** 2011. "Are Restaurants Really Supersizing America?" *American Economic Journal: Applied Economics* 3 (1): 152–88.
- Angelucci, Manuela, Silvia Prina, Heather Royer, and Anya Samek.** 2020. "Incentives and Unintended Consequences: Spillover Effects in Food Choice." *American Economic Journal: Economic Policy* 11 (4): 66–95.
- Augurzyk, Boris, Thomas K. Bauer, Arndt R. Reichert, Christoph M. Schmidt, and Harald Tauchmann.** 2018. "Habit Formation, Obesity, and Cash Rewards." Ruhr Economic Paper 750.
- Babcock, Philip, Kelly Bedard, Gary Charness, John Hartman, and Heather Royer.** 2015. "Letting Down the Team? Social Effects of Team Incentives." *Journal of the European Economic Association* 13 (5): 841–70.
- Belot, Michèle, Jonathan James, and Patrick Nolen.** 2016. "Incentives and Children's Dietary Choices: A Field Experiment in Primary Schools." *Journal of Health Economics* 50: 213–29.
- Breuer, Christoph, and Svenja Feiler.** 2015. "Sportvereine in Deutschland—ein Überblick." In *Sportentwicklungsbericht 2013/2014—Analyse zur Situation der Sportvereine in Deutschland*, edited by Christoph Breuer. Köln: Sportverlag Strauß.
- Breuer, Christoph, Svenja Feiler, and Pamela Wicker.** 2015. "Sport Clubs in Germany." In *Sport Clubs in Europe: A Cross-National Comparative Perspective*, edited by Christoph Breuer et al., 187–208. New York: Springer.
- Breuer, Christoph, Remco Hoekman, Siegfried Nagel, and Harold van der Werff, eds.** 2015. *Sport Clubs in Europe: A Cross-National Comparative Perspective*. New York: Springer.
- Carrera, Mariana, Heather Royer, Mark Stehr, and Justin Sydnor.** 2018. "Can Financial Incentives Help People Trying to Establish New Habits? Experimental Evidence with New Gym Members." *Journal of Health Economics* 58: 202–14.
- Carrera, Mariana, Heather Royer, Mark Stehr, and Justin Sydnor.** 2020. "The Structure of Health Incentives: Evidence from a Field Experiment." *Management Science* 66 (5): 1890–1908.
- Cawley, John, David Frisvold, and Chad Meyerhoefer.** 2013. "The Impact of Physical Education on Obesity among Elementary School Children." *Journal of Health Economics* 32 (4): 743–55.

- Cawley, John, Chad Meyerhoefer, and David 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: US Department of Health and Human Services.
- Childs, Emma, and Harriet de Wit.** 2014. "Regular Exercise Is Associated with Emotional Resilience to Acute Stress in Healthy Adults." *Frontiers in Physiology* 5: 161.
- Courtemanche, Charles.** 2009. "Rising Cigarette Prices and Rising Obesity: Coincidence or Unintended Consequence?" *Journal of Health Economics* 28 (4): 781–98.
- Courtemanche, Charles.** 2011. "A Silver Lining? The Connection between Gasoline Prices and Obesity." *Economic Inquiry* 49 (3): 935–57.
- Courtemanche, Charles, Garth Heutel, and Patrick McAlvanah.** 2015. "Impatience, Incentives and Obesity." *Economic Journal* 125 (582): 1–31.
- Currie, Janet, Stefano DellaVigna, Enrico Moretti, and Vikram Pathania.** 2010. "The Effect of Fast Food Restaurants on Obesity and Weight Gain." *American Economic Journal: Economic Policy* 2 (3): 32–63.
- Currie, Janet, and Firouz Gahvari.** 2008. "Transfers in Cash and In-Kind: Theory Meets the Data." *Journal of Economic Literature* 46 (2): 333–83.
- Cutler, David M., Edward L. Glaeser, and Jesse M. Shapiro.** 2003. "Why Have Americans Become More Obese?" *Journal of Economic Perspectives* 17 (3): 93–118.
- DellaVigna, Stefano, and Ulrike Malmendier.** 2006. "Paying Not to Go to the Gym." *American Economic Review* 96 (3): 694–719.
- Di Bartolomeo, Giovanni, and Stefano 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, Davide, and Nicolas 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, Pierre, Rachel Griffith, and Aviv Nevo.** 2014. "Do Prices and Attributes Explain International Differences in Food Purchases?" *American Economic Review* 104 (3): 832–67.
- Dunn, Richard 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): 1149–64.
- 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, Christina, and Rafael Lalive.** 2018. "Does Early Child Care Affect Children's Development?" *Journal of Public Economics* 159: 33–53.
- Fletcher, Jason M., David E. Frisvold, and Nathan Tefft.** 2010. "The Effects of Soft Drink Taxes on Child and Adolescent Consumption and Weight Outcomes." *Journal of Public Economics* 94 (11–12): 967–74.
- Fricke, Hans, Michael Lechner, and Andreas Steinmayr.** 2018. "The Effects of Incentives to Exercise on Student Performance in College." *Economics of Education Review* 66: 14–39.
- Frieden, Thomas R., William Dietz, and Janet Collins.** 2010. "Reducing Childhood Obesity through Policy Change: Acting Now to Prevent Obesity." *Health Affairs* 29 (3): 357–63.
- Friedrichsen, Jana, Tobias König, and Renke Schmacker.** 2018. "Social Image Concerns and Welfare Take-Up." *Journal of Public Economics* 168: 174–92.
- German Olympic Sport Confederation.** 2011. "Sportgutscheine für Sachsens Drittklässler." *Der Deutsche Olympische Sportbund-Presse News*. January 31. <https://www.dosb.de/sonderseiten/news/news-detail/news/sportgutscheine-fuer-sachsens-drittklaessler>.
- German Olympic Sport Confederation.** 2017. *Mitgliederentwicklung in Sportvereinen: 2000 bis 2015—Bestand, Veränderungen und Perspektiven*. Frankfurt am Main: Der Deutsche Olympische Sportbund.
- German Public Health Service.** 2018. "School Health Examination Data for an Anonymous County (Schulreihenuntersuchung)." Landratsamt (County administration). Unpublished data (accessed April 29, 2018).
- German Registry Offices.** 2017. "Registry Office Information." Selected registry offices. Unpublished data (accessed March 20, 2018).
- Goebel, Jan, Markus M. Grabka, Stefan Liebig, Martin Kroh, David Richter, Carsten Schröder, and Jürgen Schupp.** 2019. "The German Socio-Economic Panel (SOEP)." *Jahrbücher für Nationalökonomie und Statistik* 239 (2): 345–60.

- Goebel, Jan, Christian Krekel, Tim Tiefenbach, and Nicolas R. 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): 1137–80.
- Golsteyn, Bart H. H., Maria W. J. Jansen, Dave H. H. Van Kann, and Annelore M. C. Verhagen. 2020. "Does Stimulating Physical Activity Affect School Performance?" *Journal of Policy Analysis and Management* 39 (1): 64–95.
- Gould, Daniel, and Sarah Carson. 2008. "Life Skills Development through Sport: Current Status and Future Directions." *International Review of Sport and Exercise Psychology* 1 (1): 58–78.
- Graf, Sahara, and Michele 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 Paper 112.
- Griffith, Kevin N., and Yevgeniy Feyman. 2021. "Amplifying the Noise: The Dangers of Post Hoc Power Analyses." *Journal of Surgical Research* 259: A9–11.
- Griffith, Rachel, Rodrigo Lluberas, and Melanie Lührmann. 2016. "Gluttony and Sloth? Calories, Labor Market Activity and the Rise of Obesity." *Journal of the European Economic Association* 14 (6): 1253–86.
- Grossman, Michael, Erdal Tekin, and Roy Wada. 2014. "Food Prices and Body Fatness among Youths." *Economics and Human Biology* 12: 4–19.
- Hainmueller, Jens. 2012. "Entropy Balancing for Causal Effects: A Multivariate Reweighting Method to Produce Balanced Samples in Observational Studies." *Political Analysis* 20 (1): 25–46.
- Hallal, Pedro C., Cesar G. Victora, Mario R. Azevedo, and Jonathan C. K. Wells. 2006. "Adolescent Physical Activity and Health: A Systematic Review." *Sports Medicine* 36 (12): 1019–30.
- Imbens, Guido W., and Jeffrey M. Wooldridge. 2009. "Recent Developments in the Econometrics of Program Evaluation." *Journal of Economic Literature* 47 (1): 5–86.
- Jeitschko, Thomas D., and Rowena A. Pecchenino. 2006. "Do You Want Fries with That? An Exploration of Serving Size, Social Welfare, and Our Waistlines." *Economic Inquiry* 44 (3): 442–50.
- Jürges, Hendrik, Steffen Reinhold, and Martin Salm. 2011. "Does Schooling Affect Health Behavior? Evidence from the Educational Expansion in Western Germany." *Economics of Education Review* 30 (5): 862–72.
- Just, David R., and Joseph Price. 2013. "Using Incentives to Encourage Healthy Eating in Children." *Journal of Human Resources* 48 (4): 855–72.
- Kesternich, Iris, Bettina Siflinger, James P. Smith, and Joachim K. Winter. 2015. "Individual Behaviour as a Pathway between Early-Life Shocks and Adult Health: Evidence from Hunger Episodes in Post-War Germany." *Economic Journal* 125 (588): F372–93.
- Kreisportbund Landkreis Leipzig. 2019. KOMM! Engagier Dich. <https://www.ksb-ll.de/kinderjugendsport-leipzig/komm-in-den-sportverein-landkreis-leipzig/> (accessed May 7, 2019).
- Kromeyer-Hauschild, K., A. Moss, 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—Ursachen, Folgeerkrankungen, Therapie* 9 (3): 123–27.
- Kromeyer-Hauschild, K., M. Wabitsch, D. Kunze, F. Geller, H. C. Gieß, V. Hesse, A. von Hippel, et al. 2001. "Perzentile für den Body-Mass-Index für das Kindes- und Jugendalter unter Heranziehung Verschiedener Deutscher Stichproben." *Monatsschrift Kinderheilkunde* 149: 807–18.
- Lakdawalla, Darius, and Tomas Philipson. 2009. "The Growth of Obesity and Technological Change." *Economics and Human Biology* 7 (3): 283–93.
- Lally, Philippa, and Benjamin Gardner. 2013. "Promoting Habit Formation." *Health Psychology Review* 7 (S1): S137–58.
- Landessportbund Sachsen. 2019. "Kampagne KOMM! Engagier Dich: Engagementförderung im Sächsischen Sport." Leipzig: Landessportbund Sachsen. <https://www.sport-fuer-sachsen.de/sport-und-gesellschaft/engagementfoerderung/> (accessed May 4, 2019).
- Lechner, Michael. 2009. "Long-Run Labour Market and Health Effects of Individual Sports Activities." *Journal of Health Economics* 28 (4): 839–54.
- Loewenstein, George, Joseph Price, and Kevin Volpp. 2016. "Habit Formation in Children: Evidence from Incentives for Healthy Eating." *Journal of Health Economics* 45: 47–54.
- Marcus, Jan, Thomas Siedler, and Nicolas R. Ziebarth. 2021. "The Long-Run Effects of Sports Club Vouchers for Primary School Children." IZA Discussion Paper 14246.
- Marcus, Jan, Thomas Siedler, and Nicolas R. Ziebarth. 2022. "Replication data for: The Long-Run Effects for Sports Club Vouchers for Primary School Children." American Economic Association

- [publisher], Inter-university Consortium for Political and Social Research [distributor]. <https://doi.org/10.38886/E138922V1>.
- Piercy, Katrina L., Richard P. Troiano, Rachel M. Ballard, Susan A. Carlson, Janet E. Fulton, Deborah A. Galuska, Stephanie M. George, and Richard D. Olson.** 2018. "The Physical Activity Guidelines for Americans." *Journal of the American Medical Association* 320 (19): 2020–28.
- Prentice-Dunn, Hannah, and Steven Prentice-Dunn.** 2012. "Physical Activity, Sedentary Behavior, and Childhood Obesity: A Review of Cross-Sectional Studies." *Psychology, Health and Medicine* 17 (3): 255–73.
- Prina, Silvia, and Heather Royer.** 2014. "The Importance of Parental Knowledge: Evidence from Weight Report Cards in Mexico." *Journal of Health Economics* 37: 232–47.
- Reichert, Arndt R.** 2015. "Obesity, Weight Loss, and Employment Prospects: Evidence from a Randomized Trial." *Journal of Human Resources* 50 (3): 759–810.
- Royer, Heather, Mark Stehr, and Justin 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, Martin, and Daniel Schunk.** 2012. "The Relationship between Child Health, Developmental Gaps, and Parental Education: Evidence from Administrative Data." *Journal of the European Economic Association* 10 (6): 1425–49.
- Schüttoff, Ute, Tim Pawlowski, Paul Downward, and Michael Lechner.** 2018. "Sports Participation and Social Capital Formation during Adolescence." *Social Science Quarterly* 99 (2): 683–98.
- Schwerdt, Guido, Dolores Messer, Ludger Woessmann, and Stefan C. Wolter.** 2012. "The Impact of an Adult Education Voucher Program: Evidence from a Randomized Field Experiment." *Journal of Public Economics* 96 (7–8): 569–83.
- Siedler, Thomas, Jürgen Schupp, C. Katharina Spiess, and Gert G. Wagner.** 2009. "The German Socio-Economic Panel (SOEP) as Reference Data Set." *Schmollers Jahrbuch* 129 (2): 367–74.
- Statistisches Landesamt, Freistaat Sachsen.** 2020. "Bevölkerungsstand des Freistaates Sachsen nach Alter und Geschlecht." https://www.statistik.sachsen.de/download/100_Berichte-A/A_I_3_j18_SN.pdf (accessed March 26, 2020).
- Stevenson, Betsey.** 2010. "Beyond the Classroom: Using Title IX to Measure the Return to High School Sports." *Review of Economics and Statistics* 92 (2): 284–301.
- Strulik, Holger.** 2014. "A Mass Phenomenon: The Social Evolution of Obesity." *Journal of Health Economics* 33: 113–25.
- World Health Organization.** 2017. *Report of the Commission on Ending Childhood Obesity—Implementation Plan: Executive Summary*. Geneva: World Health Organization.
- World Health Organization.** 2018. Physical Activity. <https://www.who.int/news-room/fact-sheets/detail/physical-activity> (accessed March 26, 2020).
- World Health Organization.** 2020. "Noncommunicable Diseases: Childhood Overweight and Obesity." <https://www.who.int/dietphysicalactivity/childhood/en/> (accessed March 26, 2020).
- Züchner, Ivo, and Bettina Arnoldt.** 2012. "Sport von Kinder und Jugendlichen in Ganztagschule und Sportverein." *Körper, Bewegung und Schule* 3 (6): 1–13.