The closer the sportier?

Children's sport activity and their distance to sport facilities

Andreas Steinmayr, Christina Felfe, Michael Lechner*

Swiss Institute for Empirical Economic Research

University of St.Gallen

This draft: February 2011

Date this version was printed: ***22 February 2011***

Comments are very welcome

Abstract: We investigate whether the distance between the next sport facilities and children's homes matter for their sports activities inside and outside of sport clubs. Our analysis is based on a large and informative cross-section of individual data of children and their families, the so-called KIGGS data. We use a semiparametric econometric method to investigate this relationship empirically. Our results suggest that while the distance does not matter in larger towns and cities, it does matter in smaller towns and in particular on the countryside.

Keywords: Sport activities of children, KIGGS data, propensity score matching methods.

JEL classification: I12, H42

Address for correspondence: Christina Felfe, Michael Lechner, Andreas Steinmayr
Swiss Institute for Empirical Economic Research (SEW), University of St. Gallen, Varnbüelstrasse 14, CH-9000 St. Gallen, Switzerland, Christina.Felfe@unisg.ch, Michael.Lechner@unisg.ch
Andreas.Steinmayr@unisg.ch, www.sew.unisg.ch/lechner.

---

* Michael Lechner is also affiliated with CEPR and PSI, London, CES-Ifo, Munich, IAB, Nuremberg, IZA, Bonn, and ZEW, Mannheim. Christina Felfe is also affiliated with CES-Ifo, Munich. This project received financial support from the St. Gallen Research Center in Aging, Welfare, and Labour Market Analysis (SCALA). We are very much indebted to the Robert-Koch Institut in Berlin, in particular Dr. Martin Schlaud, for providing the data of the KIGGS study, helping us to better understand it, and to organise a process that allowed us to merge the sports facility information to the individual data. A previous version of the paper was presented at the first European conference on sports economics in Cologne, 2010. We thank participants for helpful comments. The usual disclaimer applies.

1 Introduction

The positive effect of physical activities on individual health is widely acknowledged both in academics and in the general public (Lakdawalla & Philipson, 2007; Lechner, 2009; Gomez-Pinilla, 2008; Rashad, 2007; among others). Nevertheless, a substantial part of the population is not involved in individual sports activities. In Germany, for instance, about 40% of the population older than 18 does not participate in sports activities at all, a share that is close to the European average (Bundestag, 2006; Gratton & Taylor, 2000). Among German children in the age range 6-14, this number is somewhat lower, about 30% (Kutteroff & Behrens, 2006). Nevertheless, given that sports during childhood is one of the key determinants of sports during adulthood, and in the same vein health during childhood is one of the key determinant of health during adulthood, this figure is still a matter of concern.

In light of these non-activity figures one may question the substantial public subsidies paid to the leisure sports sector in many Western countries (Gratton & Taylor, 2000). In Germany, for instance, about 0.2% of the German GDP is spent on the provision of sports-related goods and services, of which 77% goes directly to the construction and maintenance of sports facilities (Ahlert, 2004). Yet, there is only scarce evidence on the causal link between the availability of sport facilities and sports engagement among adults (Limstrand, 2008) and basically no evidence about this link among children.[1]

The objective of this paper is to investigate if an increase in the local availability of sports facilities is directly linked to an increase in children's sports engagement. The particular question we want to answer is whether children who live closer to a sports facility

[1] So far the literature has mainly concentrated on the association between the availability of outdoor spaces and people's physical activity (Boone-Heinonen, Casanova, Richardson, & Gordon-Larsen, 2010; Susan, Hastert, Yu, & Brown, 2008; Humphreys & Ruseski, 2007; Kaczynski, Potwarka, & Saelens, 2008; Huston, Evenson, Bors, & Gizlice, 2003), not however on the impact of specific sports facilities.

engage more in sports activities. When analyzing this question we distinguish between sports exercised inside and outside of sport clubs. In addition, we look at different types of sports facilities, such as gyms, sports grounds, tennis courts, and swimming baths, to gain some understanding which type of sports facility is relevant to encourage children to engage in sports.

The "German Health Interview and Examination Survey for Children and Adolescents" (henceforth KiGGS) - a comprehensive, Germany-wide, representative dataset for the age group 0-17 years - constitutes an excellent base to analyze the outlined questions (for a general description of this data set, see *www.kiggs.de)*. It provides not only detailed information on the intensity of children's physical activity, but it also distinguishes between sports activities exercised inside and outside of a club. The latter information is, however, only available for children in the age range 3-10 years, which is the reason why we restrict our analysis to this specific age group. Using self-collected information about the supply of different sports facilities in the local neighbourhood of each survey participant, we supplement this data with information on the distance between the home of each child and different types of sports facilities.

Yet, a simple comparison of children who are living closer to sports facilities versus children who are living further away would not provide us with reliable estimates for the causal link between the distance and the level of children’s physical activity. First, the availability of sports facilities might be a result of the lobbyism of the local citizens and second, parents' location choice might be based on the amenities offered by the neighbourhood. As a result, children, whose parents care more about them engaging in sports, are more likely to live in a neighbourhood with abundant supply of sports facilities. Nevertheless, the part of this endogeneity problem coming from the supply of facilities is less important for Germany: the majority of sports facilities has been constructed from 1960 to

1990, far before our period of interest, following the so-called Golden Plan (Hübner, 2003) - a major effort of the German government to extend and improve sports facilities. Hence, we are confident that the availability of local sports facilities is exogenous to any individual political efforts. Moreover, the data we have at hand provides us with a wide range of information about the child, her family background as well as the characteristics of her local neighbourhood. This allows us to address the second concern, namely that parents' location choice might be influenced by the availability of sports facilities. We can exploit this rich source of information by employing a matching estimator, which allows us to compare children which are similar in many individual, family and neighbourhood aspects except the local supply of sports facilities. Thus, we are confident that the results uncover the impact of the local availability of sports-related institutions on children's physical activity.

The remainder of the paper is structured as follows. The next section provides a brief overview of sports facilities in Germany. Section 3 describes the two datasets we use in detail. Section 4 shows first descriptive results about the link between facilities and activity levels and then discusses the matching results. Section 5 concludes.

2 Sports facilities in Germany

The so-called „Golden Plan for Health and Recreation“, established in the 1950ies (DOG, 1961), constitutes the key stone for the systematic planning of sports facilities in West Germany. As its target of providing $0.2m^2$ indoor and $3.25m^2$ outdoor spaces for each citizen required a big extension of the existing capacity, the plan caused a major construction wave for sports facilities. As a result, the supply of sports facilities increased dramatically from around 40.000 facilities in 1960 to around 140.000 in 1988 (Hübner, 2003). From the 1990ies onward, the supply remained rather stable with about 100.000 core sports facilities and 50.000 further sports institutions. After unification, East Germany followed the West German example and introduced the so-called “Golden Plan East”, which lead to an additional supply of 20.000 sports facilities in the East German states (Hübner, 2003).

Total public expenditure for the provision of sports-related goods and services amounts to 0.2% of the German GDP:[2] 77% of this amount (3.94 bio €) is used for the provision of sports-related services and 23% (0.9 bio €) is spent on administration (e.g. management, sports events). The overall relative spending levels vary, however, dramatically across states. They range from 0.14% in North-Rhine Westphalia and Schleswig-Holstein to 0.37% in Saxony-Anhalt and even 0.41% in Thuringia. Expressed in monetary terms, this means that on average 49.22€ are spent per person for the provision of sports-related goods and services, the minimum amount spent in Hamburg (11.29€/person) and the maximum amount spent in Baden-Wurttemberg (67.56€/person) (Ahlert, 2004).

3 Data

Our analysis draws upon the "German Health Interview and Examination Survey for Children and Adolescents" (henceforth KiGGS), which is a comprehensive, Germany-wide, representative interview and examination survey for the age group 0-17 years.[3] Between May 2003 and May 2006, 17,641 participants from 167 communities were examined and interviewed. One feature that makes KiGGS especially suitable for our purposes is that it provides not only exhaustive information on the children’s individual and family backgrounds, but also detailed information on the children's physical activities. In addition to this survey, we collected information on the available sports facilities as well as local

[2] To put that number into perspective, note that Germany spends on average 6.2% of GDP on education, including early childcare, Kindergarten, obligatory school system as well as higher education and research. Of course, some of this spending may be related to school based sports as well.

[3] For more information about KiGGS please refer to http://www.kiggs.de/service/english/index.html.

characteristics for (almost) all communities included in KiGGS (a detailed description of this database is given below). Based on the exact location of the sports facilities and the children included in KiGGS, for each child we determine straight-line and road distances to the closest sports facility, as well as to different types of sports facilities. Thus, the KiGGS data together with the detailed information on local conditions allow us to study the relation between children's sports participation and the availability of sports facilities in their local neighbourhood.

3.1 The sample

The information on children's physical activity is recorded by a set of questions that differ according to children's age. Our interest lies not only on participation in sports activities in general, but also in participation in sports clubs in particular. Given that the latter information is only reported for children age 3 to 10 years old, our analysis is restricted to this age range and thus to 8023 families. Due to missing information on the individual distances to available sports facilities in the local area (550 observations) and on the individual participation in a sport club or sports outside a club (475 observations), our sample is further restricted and finally includes 6998 children (and their families).[4]

Parents were asked about the frequency with which their child was performing sports in a club and additionally about the frequency with which their child was performing sports outside of a club. They could choose between 5 different categories: "never", "less than once per week", "once or twice a week", "3-5 times a week "and "almost daily". Table 1 shows that there are basically two groups of children, those that are not active in a sports club on a regular basis (47%) and those who are active at least once a week (53%). With respect to

[4] For an exact description about the sample construction please refer to Appendix C.

sports outside a club, the picture is similar: Half of the kids exercise at least once per week, while the other half do not exercise at all, or at least not outside a sports club - the distribution is, however, more even across all possible frequencies. Thus, in our analysis we focus on the impact of participating in sports inside (outside) a club on a regular basis (53% and 49%, respectively) versus not participating in a sports club (exercising in general) on a regular basis (47% and 50%, respectively).

Table 1: Frequency of participation in a sports club

	Sports in a club		Sports outside a club*	
	Observations	%	Observations	%
Daily	80	1	672	9
3-5 times/week	401	6	716	10
1-2 times/week	3,278	46	2,123	30
Rarely	436	6	2,044	29
Never	2,974	41	1,443	21

Note: *Sports outside a club is missing for 172 children. The percentages correspond to a total of 6998 children. All further analyses are based on this sample.

A summary of the individual, family background and local neighbourhood characteristics can be found in Table A.1 in Appendix A. With respect to the individual characteristics of the children, children are on average 6.6 years old, where each age group is equally well represented. About half of the children are boys (51%). Children's socio-economic background can be described as follows: The majority of the parents have a basic or intermediate school degree (among mothers 20% and 45%, respectively, and among fathers 25% and 34%, respectively). Almost half of the sample belongs to the middle class (46%), one quarter belongs to the lower class (27%) and one quarter to the upper class (26%). It is important to stress that 13% of the children are from a family with foreign background. Given that foreigners may behave differently in terms of engagement in social activities and in particular in sports activities (especially when their child is a girl), it is crucial to consider the cultural background in our estimation. One final noteworthy fact is that 22% of the mothers are overweight and 10% of them are even obese. Thus, physical activity seems to be highly relevant.

3.2 Information on sports facilities

The specific sports facilities considered are gyms, sports grounds, tennis courts, and indoor pools. According to the German Olympic Association the selected facilities should serve as a location for most sports performed by children aged 0 to 18 years.

Data on type and address of these facilities was collected from various information sources for 163 out of the 167 communities, where families belonging to KiGGS were surveyed. The main sources of information were municipality websites, where most municipalities provide registers of available facilities. Additionally, we collected data from websites of local sports associations as well as *GoogleMaps*. For municipalities where the relevant information could not be gathered online, we requested lists of available facilities from the municipal administration (and usually received them). [5]

The address data from the survey participants as well as from sports facilities were geo-coded using Microsoft MapPoint Europe 2010 in combination with IC-Tools, a MapPoint AddIn for Microsoft Excel.[6] Road distances between children's homes and the different sports facilities available in the community[7] were calculated using the same software. Linear

---

5 Information on sports facilities was not available or could not be processed for four communities.

6 For addresses where *MapPoint* was not able to determine the exact Geocode, *GoogleMaps* and *BingMaps* were used instead.

7 In order to keep the collection of data regarding the availability of facilities feasible, we collected only data in the communities where children were surveyed. For children who live close to the boarder of a municipality it might therefore be the case, that a specific facility in the neighboring municipality is closer than a facility of the same type in the residence municipality. In case a particular type of facility was not available in a municipality, we tried to collect the information on available facilities of this type in neighboring municipalities. However, this was not possible for all tennis courts and indoor pools. The number of observations for these two types of facilities is therefore lower than for gyms and sports grounds (see Table 2).

distances were calculated using the STATA module *globdist*. Finally for each type of facility the distance to the closest facility was obtained.

Table 2: Mean distance (km) to closest facilities

		Mean linear distance (km) to closest …			
		Gym	Sports ground	Tennis court*	Indoor pool*
Total		1.29	1.12	2.10	3.87
Gender:	Girl	1.31	1.12	2.15	3.92
	Boy	1.27	1.12	2.05	3.82
Age	3 - 6 years	1.30	1.13	2.09	3.95
	7 - 10 years	1.27	1.11	2.11	3.80
Nationality:	German	1.39	1.17	2.21	4.11
	Foreign	0.63	0.79	1.43	2.45
Social class:	Low	1.25	1.13	2.09	3.75
	Medium	1.41	1.14	2.29	4.27
	High	1.12	1.08	1.80	3.35
Living in a	Village	2.21	1.48	3.05	7.18
	Small town	0.79	1.02	1.73	3.04
	Medium town	0.85	0.99	1.75	2.45
	City	0.61	0.76	1.50	2.09
Part of Germany:	West	0.99	0.99	1.76	3.64
	East	1.92	1.39	3.29	4.44

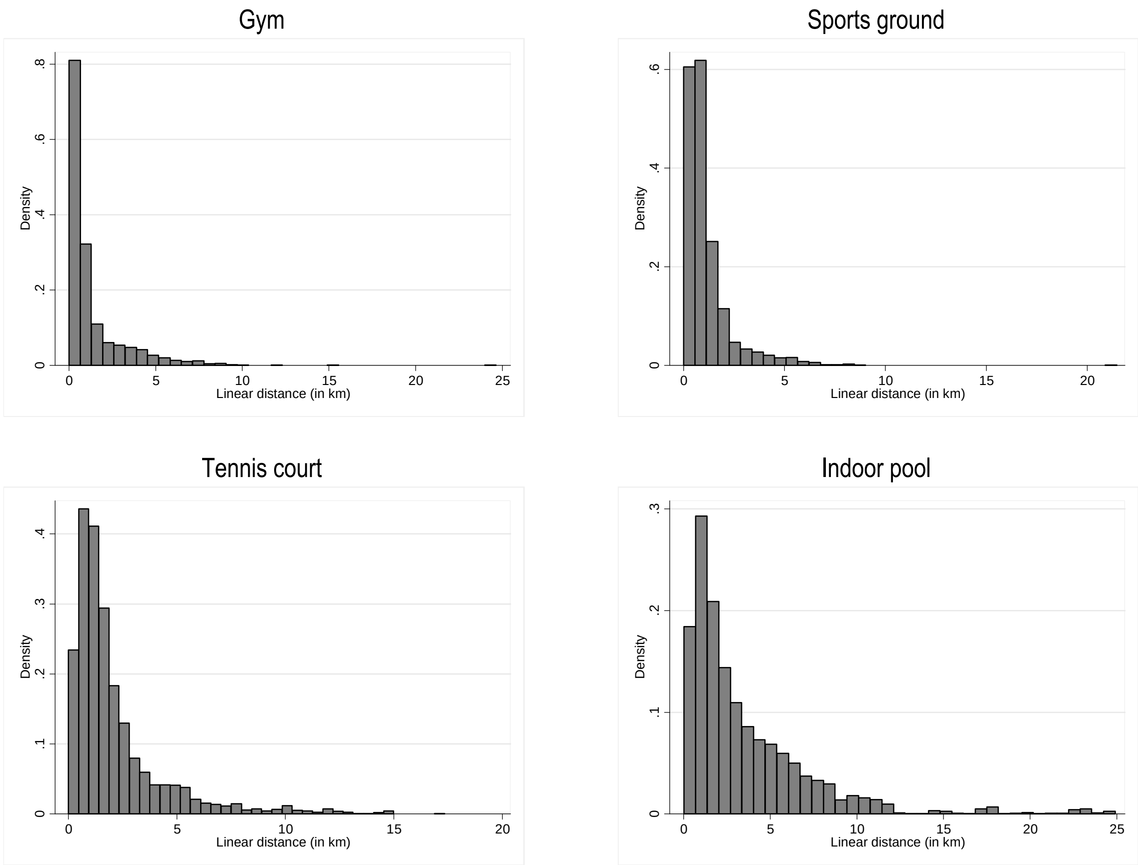
Note: Distance defined as straight line. Distance to tennis court is missing for 957 observations, distance to indoor pool for 793 observations.

Table 2 provides an overview over the mean linear distances to the closest facility of each type. On average the nearest gym is 1.29 km, the nearest sports ground 1.12 km, the nearest tennis court 2.1 km, and the nearest indoor pool 3.87 km away. While the distances are almost the same by the child's sex and age, children with an immigrant background live on average significantly closer to all types of facilities, which can be mainly attributed to the fact that immigrant families are more likely to live in cities. Due to a higher density of sports facilities within cities, the distance is necessarily shorter: while, for instance, the average distance to the closest gym is only 610 meters in cities, it is more than 2 km in villages.[8] Substantial differences exist as well between East- and West Germany.

[8] Notice that the definition of village, small town, medium town and city is based on INKAR and is a combination of population size, density, political and administrative relevance, etc. For a more detailed description please refer to http://www.bbsr.bund.de

Not surprisingly, in our sample the distribution of the distances to the four different categories of sports facilities is not at all homogeneous (see Figure 1). Many observations are clustered at lower distances (notice that many children live in larger cities), while less observations are located at larger distances beyond, for example, 2.5 km. Moreover, the distances to gyms and sports grounds are shorter than for tennis courts and indoor pools, a fact which is due to the greater availability of the first group of facilities.

Figure 1: Distribution of observations with respect to distance to next facility



Note: Distance is measured as straight line.

Table 3 shows the correlation between local availability of sports facilities and further local conditions. The negative correlation between population density, population share below 18 years and population development, on the one hand, and distance to the different sports facilities, on the other hand, as well as the positive relationship between open space per habitant and the distance to sports facilities result from the fact that there are generally less sports facilities available in smaller villages/cities than in bigger cities. The positive association between the availability of sports facilities and economic conditions (higher tax income, lower unemployment and more employment in the tertiary sector), might also result from the differential availability of sports facilities across differentially big municipalities.

Table 3: Correlation between distance to facilities and municipality characteristics

	Gym	Sports ground	Tennis court	Indoor pool
Population density	-0.30	-0.20	-0.23	-0.31
Open space per inhabitant	0.41	0.27	0.47	0.59
Population share below 18 years	-0.13	-0.05	-0.11	0.04
Tax income per capita	-0.27	-0.17	-0.28	-0.27
Unemployment rate	0.10	0.08	0.13	-0.08
Labor force in tertiary sector (%)	-0.20	-0.06	-0.17	-0.18
Population development	-0.17	-0.09	-0.24	-0.21

Note: All correlations are significant at the 1% significance level. Distance is measured as straight line.

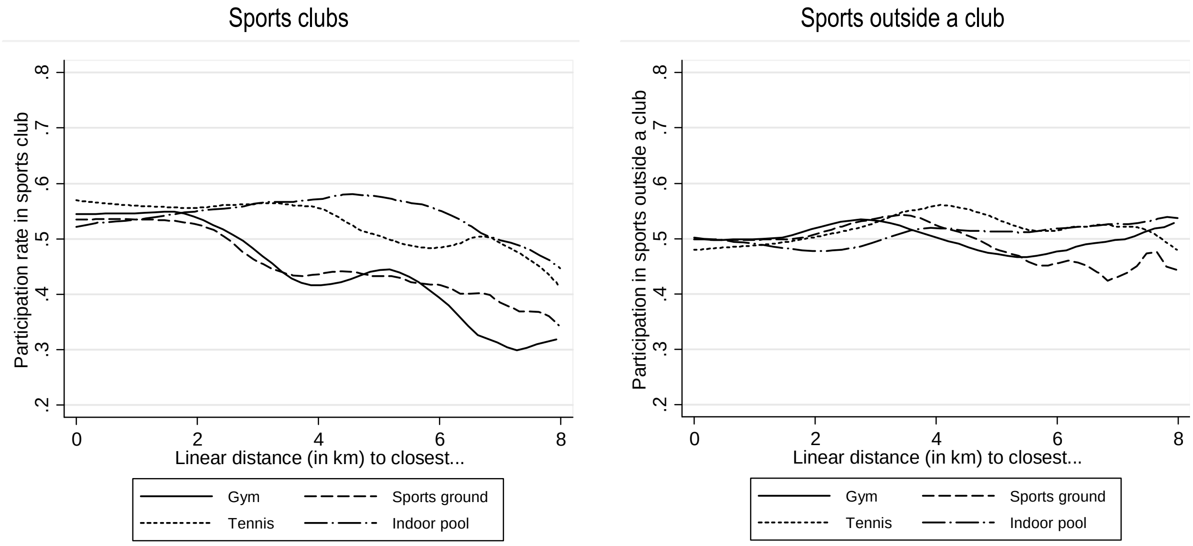
Taken together, the significant difference in the supply of sports facilities across differential family background and regions highlights the need to condition on these background characteristics in our analysis.

4 Distance to sports facilities and children's physical activity

4.1 Unconditional results

The question posed in this study is whether children who live closer to sports facilities exercise more sports. In order to gain a better understanding of how children’s physical activity evolves with the distance to different types of sports facilities, Figure 2 presents non-parametric kernel estimates of children’s sports participation - inside and outside of a club - in relation to the distance to four sports facilities - sports halls ('gyms'), sports grounds, tennis courts and indoor pools. Distance is measured as a straight line between home address and address of the sports facility. Measuring distance using roads, the findings do not change in any substantial manner (see Figure A.1 in Appendix A).

Figure 2: The relation of the distance to sports facilities to the level of sports participation



Note: Nonparametric kernel regression based on the Epanechnikov kernel. A bandwidth of 0.6 is used for both figures. Distance is measured as straight line. Figure A.1 provides the same relationships using road distances.

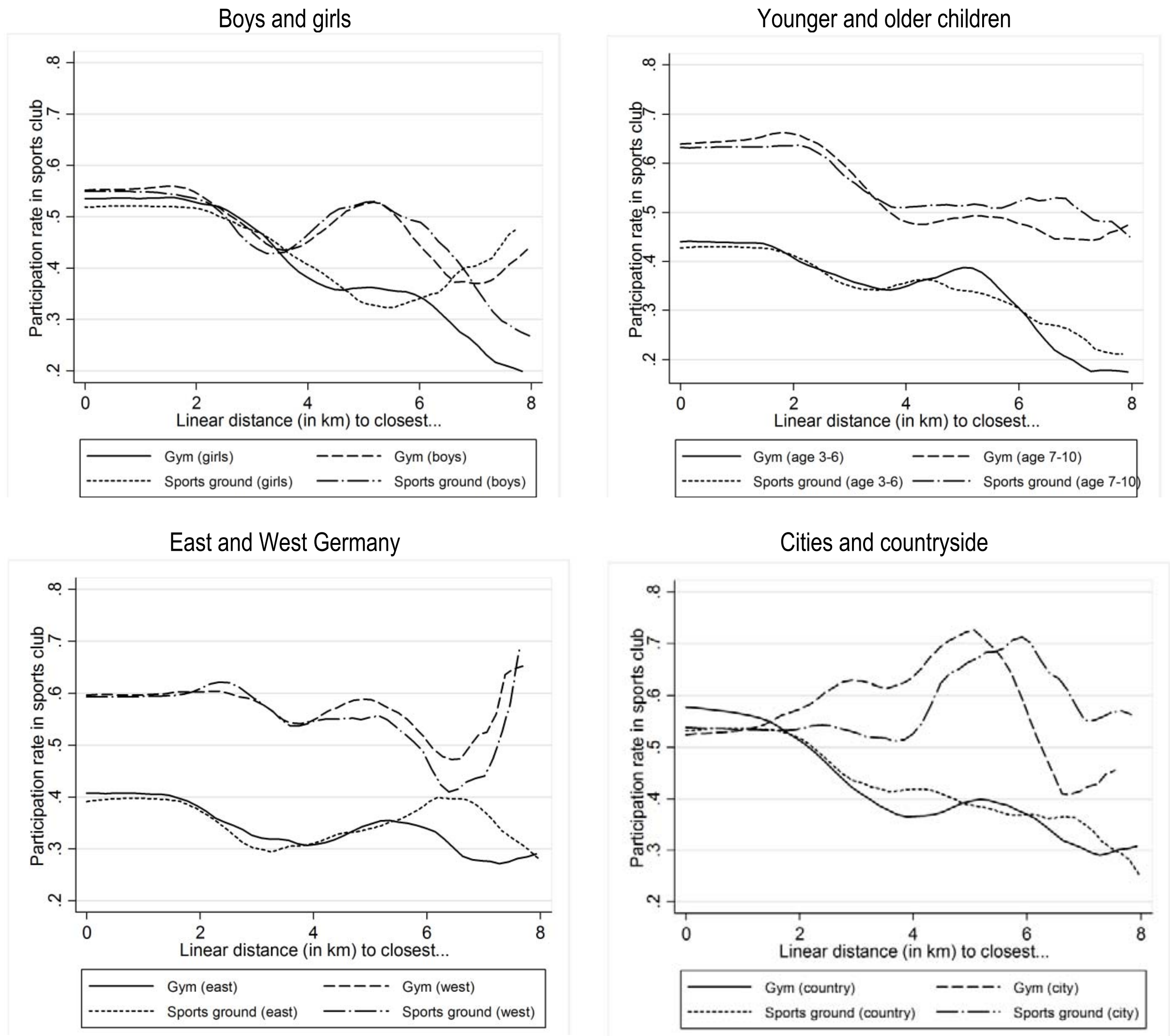
As we can see clearly in Figure 2, children's participation in sports clubs decreases the further they live from the next sports facility. This negative relation is most prominent for sports halls and sports grounds - a fact which is possibly explained by gyms and sports grounds serving as a location for most sports exercised by children. Remarkably, sports club participation is rather inelastic with respect to the local supply of sports halls/grounds over the first 2 km. The negative relation is less pronounced, but still observable with respect to the distance to the closest tennis facility or indoor pool. Yet, in the case of tennis courts children’s club participation only starts decreasing when the distance exceeds 4 km, and in the case of indoor pools even only at a distance of 6 km and more (as can be seen in Figure 1, our measurements in these areas are, however, based on very few observations).

Quite the contrary to sports exercised inside a club, children's physical activity outside a club seems to be completely inelastic to the availability of sports facilities, no matter which type of facility. Yet, given that sports not organized in clubs may be exercised independently of a sports facility - think, for instance, about running, skating, etc. - this non-relation is perhaps not surprising.

Given these results, from now on we concentrate on the impact of the local availability of sports grounds and sports halls only. In a next step, we investigate if the above presented shape holds also for specific subpopulations. For this purpose, we repeat the same exercise as above for different strata: children living on the countryside (versus children living in bigger cities), children residing in East Germany versus children residing in West Germany, boys versus girls, and last Kindergarten children (aged 3 to 6) versus school children up to the age of 10 (see Figure 2).

Figure 3 shows that boys and girls react rather similarly to the distance to the next sports hall/ground, with the exception that boys strangely increase their sports participation when living between 4 and 6 km to the next sports facility after having initially decreased their participation at km 2. When comparing children 3 to 6 years old with children 7 to 10 years old we only observe differences in the level of sports participation, older children exercise much more, but there are not much differences in the pattern of how children's sports participation changes with the distance to the next sports hall or ground. In the same vein, there are notable differences with respect to the level of the physical activity of children living in East and West Germany, not however with respect to the correlation between the physical activity and the local availability of sports halls/grounds.

Figure 3: Heterogeneity in the relation of the distance to sports facilities to the level of sports participation



Note: See note below Figure 2.

The picture changes, however, dramatically when comparing children who live in more or less urban areas. While children living in the countryside exhibit the same pattern as all other subgroups so far, there is no systematic pattern observable for children who live in larger cities. Notice that this result is confirmed by a parametric analysis of children's sports participation where we control for individual background characteristics in a probit model (see Table A.3 in Appendix A). There are several possible explanations for this rather erratic picture for children living in a city: (1) the available distance measure might not correspond to the exact walking distance – for instance, there may be shortcuts for pedestrians/cyclists; (2) public transport may allow reaching further location more or less easy; (3) the density of sports facilities may be too high in cities to observe any systematic pattern.

Given these caveats for the case of urban areas, our analysis focuses from now on exclusively on children living in less urban areas. Thus, the following estimation results are based on a sample of 3404 observations living in the 'country side', which includes small towns and villages.

4.2 Matching results

Below we provide estimates for the impact of the local availability of sports halls and sports ground on children's sports engagement in a club. For this purpose, we employ a propensity score matching estimator. This estimation method compares the physical activity of children who live close to a sports facility (treated group) to the physical activity of children who live further away from next sports facility (control group). To avoid any selection into treatment – in other words, to eliminate any bias which may arise due to the fact that parents who want their children to engage in sports may chose to live closer to a sports facility – we adjust the two groups for differences in their covariate composition (for a discussion about the covariates please refer to Section 3). The adjustment is based on the propensity to live close to a sports facility, which is predicted using a probit estimation where the dependent variable is a binary indicator for having a sports facility close by (the exact list of covariates used and the estimation results are shown in Table A.2 in the Appendix A).[9] Note that these semi-parametric matching-type estimators have the substantial advantage compared for example to standard linear or non-linear regression types methods using

[9] For more details about the estimation method, please refer to Appendix B.

parametric models that they do not require to specify the relationship between the outcome variable, the control variables and the distance. Furthermore and of course related to this, they allow the individual effects of the distance on the sports activity to be fully heterogeneous.

The remaining question in this binary setting is how to define 'living close to a sports facility'.[10] The selection of our threshold is based on the results of the non-parametric and parametric analyses of children's propensity to engage in sports activities presented in Section 4.1. Additionally, as it is clearly visible in Table A.3 in the Appendix A, the share of children being member in a sports club remains stable over the first 2.5 km and starts decreasing thereafter. This result holds true for both sports halls and sports grounds, as well as unconditional and conditional on a set of individual background variables. Given this insight, we define the binary treatment variable equal to one when living closer than 2.5 km to the next sports hall/ground and equal to zero when living further than 2.5 km. Thus, strictly speaking, we evaluate the impact of living closer than 2.5 km to the next facility on children's sports engagement. In particular, we estimate the following three different effects: (1) the average treatment effect (henceforth ATE), which measures the average effect on children's sports participation of having a sports facility close by; (2) the average treatment effect on the treated (ATET), which is the average effect for children who live closer than 2.5 km to the next sports facility; and (3) the average treatment effect for the non-treated (ATENT), which is the average effect for children who live further away than 2.5 km from the next sports facility. Table 4 displays the three effects of living close to a sports hall or sports ground on children's sports participation in a club.[11]

---

[10] The binary setting can easily be extended to allow for the effect of various differences in distances. However, Figure 2 strongly suggests there exist only two groups and thus, such extension would not lead to any relevant gain.

[11] Results for the impact of living close to a sports hall/ground on children's physical activities outside a club are shown in Table A.5 in the Appendix A.

In line with the unconditional results, presented in Section 4.1., children who have a sports hall close by are more likely to join a sports club (see Table 4, Panel A). The average impact amounts to 13 percentage points (significant at the 1% level), which given that only 41 % of the control group (see Table 4, Column 1) are member of a sports club, is a non-negligible effect. Distinguishing between the effect for the children living close and those living further away, reveals that this effect is persistent and quite similar in all groups. A child who lives now close to a sports hall would be 12 percentage points (significant at the 1% level) less likely to join a sports club if this hall would be removed, while a child who lives now far from a sports hall would be 14 percentage points (significant at the 5% level) more likely to join a sports club if a new hall would open up in her vicinity.

Table 4: Effect of living close to a sports facility on participation in a sports club

	Y_0 (ATE)	ATE		ATET		ATENT		No of observations	Common support	Share in common support
		θ	p-val. %	Θ	p-val. %	θ	p-val. %			
Panel A										
living close (<2.5km) to…										
Gym	0.41	0.13	*0*	0.12	*1*	0.14	*2*	3404	3137	92.2
Sports ground	0.40	0.09	*2*	0.10	*2*	0.04	*32*	3404	3004	88.2
Panel B										
Effect heterogeneity (living close to gym)										
Boys	0.45	0.03	*72*	0.05	*42*	0.00	*100*	1725	891	51.7
Girls	0.30	0.17	*0*	0.19	*0*	0.16	*1*	1679	907	54.0
3-6 years old	0.25	0.16	*2*	0.22	*0*	0.04	*41*	1688	1473	87.3
7-10 years old	0.48	0.08	*3*	0.08	*5*	0.07	*13*	1716	911	53.1
East Germany	0.30	0.06	*6*	0.07	*17*	0.06	*17*	1362	1020	74.9
West Germany	0.48	0.16	*0*	0.18	*0*	0.12	*3*	2042	734	35.9
Panel C										
Effect heterogeneity (living close to sports ground)										
Boys	0.47	0.05	*41*	0.06	*38*	0.01	*93*	1725	1499	86.9
Girls	0.33	0.14	*1*	0.15	*1*	0.08	*5*	1679	1450	86.4
3-6 years old	0.24	0.16	*1*	0.17	*0*	0.09	*10*	1688	1470	87.1
7-10 years old	0.49	0.04	*45*	0.05	*40*	0.02	*87*	1716	603	35.1
East Germany	0.27	0.08	*6*	0.09	*5*	0.01	*79*	1362	1328	97.5
West Germany	0.41	0.21	*0*	0.22	*0*	0.15	*0*	2042	1616	79.1

Note: p-values obtained from 999 bootstrap replications.

The results are similar for the sports ground (see Panel A, Table 4), but a bit lower: the overall effect amounts to 9 percentage points, the effect for the treated children to 10 percentage points (both significant at the 5% level), only the effect for the non-treated children is slightly smaller (and insignificant).

Notice that living close to a sports facility, independently of the type of the sports facility, does not have a significant impact on children's physical activity outside of a sports club (see Table A.5, Panel A) though - a finding which is again in line with the unconditional results.

Do these average effects differ across different groups of the population? In order to address this question we again divide our sample into different subgroups - boys, girls, younger children (age 3 to 6 years old), older children (age 7 to 10 years old), East and West German children - and estimate the different effects separately for each subgroup. Results are shown in Table 4, Panel B, for the treatment "living close to a gym" and in Table 4, Panel C, for the treatment "living close to a sports ground".[12]

Girls apparently benefit more from having a sports hall in their neighbourhood: the average effect for girls amounts to 17 percentage points (significant at the 1 % level), while the average effect for boys is rather small (3 percentage points) and insignificant.[13] This

---

[12] Results for the effect heterogeneity with respect to sports exercised outside a club are shown in Panel B and C in Table A3.

[13] Notice that we restrict our sample only to observations that fulfill the common support - in other words, estimation is only performed with observations that have comparable counterparts in the other treatment state. Given the sampling procedure of the KiGGs survey - 17641 children from 167 municipalities were interviewed - and the fact that several regional variables, such as the population density or the recreation area, are strong predictors of the distance to the next sports facility (see Appendix A), ensuring common support is likely to reduce the sample significantly (see for instance the common support for girls and boys in Table 4). Thus, our estimates for the subgroups (which use the same specification of the propensity but of course have lower decrease of freedom and thus higher in-sample predictive power), despite being informative, may lack precision. Therefore, we are able to discuss whether one group may experience a stronger treatment effect than the other, but we cannot necessarily judge if they are significantly different from each other.

differentially strong impact on girls and boys remains when looking at the child living close by or further away: girls who have a gym in their vicinity would reduce their physical activity by 19 percentage points if the gym were removed, while girls who live further away from the next gym would increase their physical activity by 16 percentage points if a gym were built in their vicinity. Given the initially lower level of physical activity for girls - 30% versus 45% (see Table 4, column 1) - one explanation for this finding may be that many of the boys exercise no matter how far the next sport facility is located (and the remaining boys are hard to be induced to exercise anyway), while girls may be discouraged to exercise if they have first to cross a substantial distance. But this is, of course, purely speculative and not supported by our empirical evidence.

There are also remarkable differences between younger and older children. While for children at school age the distance to the next gym does not much influence their sports engagement, for children at Kindergarten age the distance to the next gym matters. The average treatment effect corresponds to 16 percentage points, and the average treatment effect for the treated amounts to even 22 percentage points (both significant at the 1% level). The difference between younger and older children may perhaps be explained as follows: (1) children 3 to 6 years old depend fully on their parents and thus, living further away from a gym means that their parents have to spend a significant amount of time to bring their children there; (2) children age 7 to 10 years old may be able to exercise sports in a gym either at school or close by the school and, thus, it may be rather the distance from the school than the distance from home which matters for their club membership.

Lastly, the availability of sports facilities influences children's sports club participation much stronger in West Germany than in East Germany: while West German children are on average 16 percentage points more likely to join a club when there is a gym close by, East German children increase their club activity by only 6 percentage points when they live close to a gym.[14] The strong difference between East and West German children remains when comparing the effects for the treated and non-treated children (7 versus 18 percentage points and 6 versus 12 percentage points, respectively). One explanation may be that in the former GDR there was no real culture to perform sports in sports clubs and thus, even if the existence of a sports hall in the neighbourhood may not encourage East German children to join a sports club. As a consequence, even an increased effort to open up more gyms in East Germany does not seem to lead to a closure of the notable difference in the sports engagement between East and West German children (30 versus 48 %).

The results shown in Table 4, Panel C, allow us to analyze as well if there is any heterogeneity in the effect of living close to a *sports ground* across the different subgroups. The picture is similar to the picture just provided for living close to a gym. Girls react much more to the availability of a sports ground than boys, so do younger children as well as children in West Germany. Explanations for these findings are probably similar to the ones provided above for the gyms. Yet, in the end, they remain speculations.

5 Conclusions

In this paper, we analyze the impact of the distance between children's home and various sport facilities on the sports activity level of children who are 3 to 10 years old. We base this analysis on very informative individual data from the "German Health Interview and Examination Survey for Children and Adolescents" (KiGGS) - a comprehensive, Germany-wide, representative dataset. Our main empirical results stem from robust econometric methods based on optimised propensity score matching estimators. They suggest that

[14] It is important to point out that given the strong predictive power of the population density the common support condition among West German children is only for 36% fulfilled.

differences in distances to facilities do not matter much in larger cities, which is probably not surprising given the high density of facilities in such locations. However, in smaller cities and villages, they matter substantially. Moving closer to a facility may easily increase a child's likelihood to participate in some sports (organized by sports club) by more than 10 percentage points. We also find interesting heterogeneity across these effects: The effects are considerably larger for girls than for boys, for younger children than for older children, and for children living in West Germany than for children in East Germany. However, these effects may be valid only for sports exercised in sports clubs and not for sports done outside of clubs.

Giving that one may safely conjecture that sports activities improve the health of the children, increasing the number of facilities (in the vicinity of children’s residence!) appears to be an investment that improves public health (although, of course, our results are silent about its cost-effectiveness).

Our study can be improved in many dimensions, though. For example, currently the KiGGS is a cross-section and thus, we were not able to exploit or take care of any dynamics which are inherent to this problem. A panel data set could not only help to improve the identification of the estimated effects but could also help to uncover some interesting dynamics in physical activities of children when they age. Obviously, one desires a more precise and detailed measurement of the type of activities jointly with a larger sample that enables investigating the heterogeneity in a more robust way. Other obvious extensions concern the effects for children older than 10 years. Finally, one may want to use the employed distance measures to obtain better identification of the effect of sports activity of children's cognitive abilities for example by using some instrumental variable strategy. Indeed, this is the thrust of a current project of us (Felfe, Lechner, & Steinmayr, 2011).

Literaturverzeichnis

Ahlert, G. (2004). Investive Sportförderung in der Bundesrepublik. *GWS Discussion Paper* .

Boone-Heinonen, J., Casanova, K., Richardson, A., & Gordon-Larsen, P. (2010). Where can they play? Outdoor spaces and physical activity among adolescents in U.S. urbanized areas. *Preventive Medicine (51)* , 295-8.

Bundestag, D. (2006). *11. Sportbericht der Bundesregierung.* Berlin: Drucksache des Deutschen Bundestags, 16/3750.

DOG. (1961). *Der Goldene Plan in den Gemeinden. Ein Handbuch.* Frankfurt am Main: Deutsche Olympische Gesellschaft.

Felfe, C., Lechner, M., & Steinmayr, A. (2011). Join the club! Physical Activity and Child Development. *mimo* .

Gomez-Pinilla, F. (2008). The influences of diet and exercise on mental health through hormensis. *Aging Research Review, 7* , 49-62.

Gratton, C., & Taylor, P. (2000). *The Economics of Sport and Recreation.* London: Taylor and Francis.

Hübner, H. (2003). *Sportstättenentwicklungsplanung in Deutschland –Notizen zur gegenwärtigen Situation.* dvs-Informationen 18.

Humphreys, B., & Ruseski, J. (2007). Participation in Physical Activity and Government Spending on Parks and Recreation. *Contemporary Economic Policy 25(4)* , 538-552.

Huston, S., Evenson, K., Bors, P., & Gizlice, Z. (2003). Neighborhood Environment, Access to Places for Activity, and Leisure-time Physical Activity in a Diverse North Carolina Population. *American Journal of Health Promotion 18(1)* , 58-69.

Kaczynski, A., Potwarka, L., & Saelens, B. (2008). Association of Park Size, Distance, and Features with Physical Activity in Neighborhood Parks. *American Journal of Public Health 98(8)* , 1451-1456.

Kutteroff, A., & Behrens, P. (2006). *KIM Studie 2006: Kinder und Medien, Computer und Internet.* Medienpädagogischer Forschungsverbund Südwest.

Lakdawalla, D., & Philipson, T. (2007). Labor Supply and Weight. *Journal of Human Resources 42* , 85–116.

Lechner, M. (2009). Long-Run Labour Market Effects of Individual Sports Activities. *Journal of Health Economics 28(4)* , 839-854.

Limstrand, T. (2008). Environmental Characteristics Relevant to Young People's use of Sports Facilities: a Review. *Scandinavian Journal of Medicine & Science 18(3)* , 275-87.

Rashad, I. (2007). Cycling: An Increasingly Untouched Source of Physical and Mental Health. *NBER Working Paper 12929* .

Susan, B., Hastert, T., Yu, H., & Brown, R. (2008). Physical Activity Among Adolescents: When Do Parks Matter? *American Journal of Preventive Medicine 34(4)* , 345–348.

Appendix A: Further descriptive statistics

Table A.1: Descriptive statistics of the control variables

		Total		Distance to gym <= 2.5 km		Distance to gym > 2.5 km	
		Mean	Std. dev.	Mean	Std. dev.	Mean	Std. dev.
Children's characteristics							
Age (in years)		6.57	2.27	6.58	2.27	6.52	2.31
Gender (1=male, 0=female)		0.51		0.51		0.50	
German background		0.87		0.85		0.97	
Parents' characteristics							
Single parent household		0.10		0.11		0.07	
Mother's education	Basic	0.20		0.21		0.16	
	Intermediate	0.45		0.42		0.57	
	High school	0.15		0.16		0.12	
	University	0.16		0.16		0.13	
	Other	0.04		0.04		0.02	
Father's education	Basic	0.25		0.25		0.24	
	Intermediate	0.34		0.31		0.47	
	High school	0.10		0.11		0.07	
	University	0.22		0.23		0.15	
	Other	0.04		0.04		0.02	
Mother's BMI	Underweight	0.03		0.03		0.02	
	Normal	0.63		0.63		0.62	
	Overweight	0.22		0.22		0.24	
	Obese	0.11		0.11		0.11	
Social class	Low	0.27		0.27		0.27	
	Medium	0.46		0.45		0.52	
	High	0.26		0.28		0.20	
Smoking during pregnancy	Regularly	0.05		0.05		0.03	
	Occassionally	0.12		0.13		0.09	
	Never	0.82		0.81		0.86	
Neighborhood characteristics							
East Germany		0.33		0.28		0.59	
Municipality size	Village	0.36		0.28		0.79	
	Small town	0.12		0.14		0.04	
	Medium town	0.31		0.34		0.14	
	City	0.21		0.24		0.03	
Population density (inhabitants/km^{2})		733.19	887.10	838.24	918.80	184.75	360.40
Recreation area/Capita (m^{2})		39.99	37.44	36.50	30.51	58.23	58.86
Tax income/Capita (Euro)		551.80	251.63	579.59	252.10	406.76	192.49
Share of population <18 years		16.92	3.14	17.12	3.12	15.89	3.06

Note: Only variables included that are used in one of the estimation steps.

Table A.2: Results of propensity score estimation for distance to next gym and distance to next sports ground: Country sample

		Gym		Sports ground	
		Coef.	p-val. %	Coef.	p-val. %
Constant		-0.75	19.5	-1.69	0.7
Gender (1=male, 0=female)		0.05	36.1	-0.03	65.8
Age in years:	3	-0.02	81.5	-0.23	5.4
	4	0.02	87.3	-0.11	36.8
	5	0.06	54.9	-0.09	44.5
	6 (reference)				
	7	-0.04	67.7	-0.14	25.3
	8	0.03	76.6	-0.07	55.7
	9	-0.06	59.8	-0.19	11.3
	10	-0.03	75.3	-0.22	7.3
German background		-0.48	0.1	-0.66	0.1
Mother's education:	Basic	0.00	98.4	0.07	45.1
	Intermediate (reference)				
	High school	-0.16	7.7	0.08	42.8
	University	0.05	61.5	-0.10	37.3
Father's education:	Basic	-0.16	3.5	-0.20	1.7
	Intermediate (reference)				
	High school	0.04	69.4	-0.15	19.4
	University	0.11	34.3	-0.19	12.0
Mother's BMI:	Underweight or Normal (reference)				
	Overweight	0.02	74.0	0.07	34.9
	Obese	0.00	98.2	0.01	95.5
East Germany x Social Class:	Low	-0.12	20.4	-0.21	4.3
	Medium (reference)				
	High	-0.08	52.1	-0.01	92.6
West Germany x Social Class:	Low	-0.10	33.2	-0.15	18.0
	Medium (reference)				
	High	0.04	75.3	0.20	17.7
Single parent household		0.29	0.5	0.22	6.0
Smoking during pregnancy: Never		-0.06	43.5	0.02	84.8
East Germany		-0.40	53.5	2.10	0.4
Municipality size:	Village	-0.78	0.0	-0.26	4.2
	Small towns (reference)				
East Germany x Population density (=Low)		-0.33	1.5	-0.25	11.8
	(=Medium or High) (reference)				

Table A.2 to be continued …

Table A.2 … continued

	Gym		Sports ground	
	coef.	p-val. %	coef.	p-val. %
West Germany x (Population density=Low)	0.12	30.7	0.24	8.0
(=Medium or High) (reference)				
East Germany x Recreation area(=Low)	-0.71	0.0	0.38	0.5
(=Medium) (reference)				
(=High)	-0.10	44.0	0.35	0.3
West Germany x Recreation area(=Low)	0.51	0.0	0.23	4.6
(=Medium) (reference)				
(=High)	0.32	0.1	0.04	69.5
East Germany x Tax income(=Low)	0.76	0.0	0.76	0.0
(=Medium) (reference)				
(=High)	-0.64	0.0	0.04	75.8
West Germany x Tax income(=Low)	-0.26	0.4	-0.01	88.7
(=Medium) (reference)				
(=High)	-0.18	12.2	-0.35	0.5
East Germany x Population share below 18 years(=Low)	0.38	0.0	0.02	82.4
(=Medium) (reference)				
(=High)	-0.04	68.5	-0.18	9.6
West Germany x Population share below 18 years(=Low)	-0.25	2.9	0.22	9.0
(=Medium) (reference)				
(=High)	0.43	0.0	0.31	0.1
East Germany x log(Population Density)	0.58	0.0	0.26	0.6
West Germany x log(Population Density)	0.57	0.0	0.78	0.0

Note: Probit model. Dependent variable is 'distance less than 2.5 km. Efron's Pseudo R^2 is 0.32 for gyms and 0.13 for sport grounds.

Table A.3: The dependence of sports participation rate on the distances conditional on covariates: Country sample

	No. of observations	Sports in club			Sports not in club		
		Unconditional	Conditional	Std. Dev.	Unconditional	Conditional	Std. Dev.
Gym							
0 - 0.5 km	1005	0.59	0.56	0.02	0.51	0.53	0.02
0.5 - 1 km	802	0.55	0.52	0.02	0.50	0.50	0.02
1 - 1.5 km	340	0.53	0.50	0.03	0.53	0.52	0.03
1.5 - 2 km	197	0.50	0.49	0.04	0.52	0.52	0.04
2 - 2.5 km	129	0.54	0.53	0.04	0.59	0.58	0.04
2.5 - 3 km	132	0.42	0.43	0.04	0.52	0.52	0.04
3 - 4 km	267	0.32	0.38	0.03	0.54	0.55	0.03
4 - 5 km	220	0.39	0.45	0.03	0.46	0.46	0.04
5 + km	312	0.36	0.47	0.03	0.48	0.46	0.03
Sports ground							
0 - 0.5 km	957	0.52	0.52	0.02	0.52	0.52	0.02
0.5 - 1 km	960	0.52	0.50	0.02	0.50	0.50	0.02
1 - 1.5 km	533	0.57	0.53	0.02	0.50	0.49	0.02
1.5 - 2 km	284	0.49	0.52	0.03	0.51	0.51	0.03
2 - 2.5 km	178	0.47	0.52	0.04	0.53	0.55	0.04
2.5 - 3 km	77	0.35	0.35	0.05	0.56	0.53	0.06
3 - 4 km	180	0.42	0.44	0.04	0.53	0.51	0.04
4 - 5 km	105	0.42	0.48	0.04	0.50	0.50	0.05
5 + km	130	0.34	0.46	0.04	0.48	0.49	0.05

Note: Results from probit estimations with the respective sports activities as dependent variables. Dummies for distances (shown) and the control variables of the propensity score are included as explanatory variables. Average discrete effects of the distance dummies and their standard error are shown. Standard errors obtained from 999 bootstrap replications. Efron's Pseudo R^2 is 0.14 for sports in clubs using distance to gym and 0.14 for distance to sports ground. For sports outside clubs and distance to gym as well as distance to sports ground Efron's Pseudo R^2 is 0.03.

Table A.4: The dependence of sports participation on the distances conditional on covariates: City sample

	No. of observations	Sports in club			Sports not in club		
		Uncond.	Conditional	Std. Dev.	Uncond.	Conditional	Std. Dev.
Gym							
0 - 0.5 km	1911	0.50	0.53	0.01	0.49	0.50	0.01
0.5 - 1 km	1028	0.55	0.53	0.01	0.50	0.49	0.02
1 - 1.5 km	264	0.50	0.47	0.03	0.45	0.44	0.03
1.5 - 2 km	112	0.57	0.52	0.04	0.57	0.57	0.05
2 - 2.5 km	85	0.68	0.60	0.05	0.48	0.47	0.06
2.5 - 3 km	51	0.65	0.60	0.06	0.57	0.56	0.07
3 - 4 km	80	0.58	0.56	0.05	0.46	0.47	0.05
4 - 5 km	36	0.75	0.69	0.07	0.50	0.51	0.08
5 + km	27	0.52	0.47	0.09	0.59	0.58	0.10
Sports ground							
0 - 0.5 km	1084	0.54	0.54	0.01	0.49	0.50	0.02
0.5 - 1 km	1405	0.52	0.52	0.01	0.48	0.48	0.01
1 - 1.5 km	687	0.53	0.53	0.02	0.51	0.50	0.02
1.5 - 2 km	229	0.55	0.57	0.03	0.50	0.50	0.03
2 - 2.5 km	71	0.54	0.49	0.05	0.51	0.51	0.06
2.5 - 3 km	40	0.58	0.53	0.07	0.55	0.55	0.08
3 - 4 km	36	0.39	0.37	0.07	0.53	0.52	0.08
4 - 5 km	20	0.65	0.58	0.10	0.60	0.59	0.12
5 + km	22	0.64	0.60	0.11	0.50	0.51	0.11

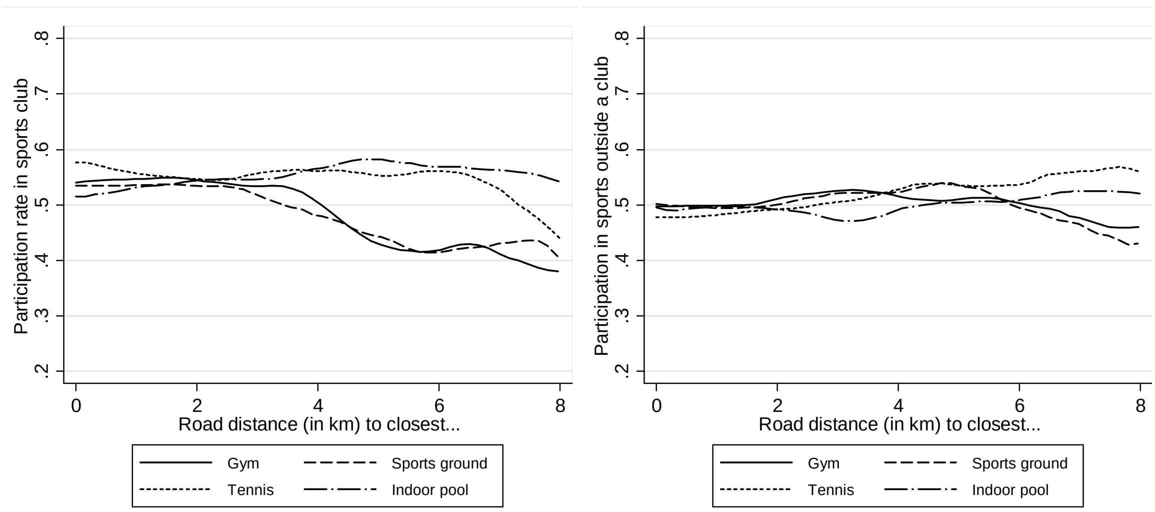
Note: See note below Table A.3. Efron's Pseudo R^2 is 0.16 for sports in clubs using distance to gym and 0.16 for distance to sports ground. For sports outside clubs and distance to gym Efron's Pseudo R^2 is 0.03 and for distance to sports ground 0.02.

Table A.5: Effect of living close to a sports facility on participation in sports outside a club

	Y_0 (ATE)	ATE		ATET		ATENT		No. of observations	Common Support	Share in common support
		θ	p-val. %	θ	p-val. %	Θ	p-val. %			
Panel A										
living close (<2.5km) to…										
Gym	0.46	0.08	*18*	0.08	*27*	0.06	*16*	3404	3137	92.2
Sports ground	0.52	0.01	*78*	0.00	*99*	0.06	*24*	3404	3004	88.2
Panel B										
Effect heterogeneity (living close to gym)										
Boys	0.56	0.03	*50*	-0.04	*58*	0.11	*12*	1725	891	51.7
Girls	0.54	-0.04	*41*	-0.10	*23*	0.03	*55*	1679	907	54.0
3-6 years old	0.41	0.11	*16*	0.11	*33*	0.11	*4*	1688	1473	87.3
7-10 years old	0.56	-0.01	*77*	-0.04	*39*	0.02	*65*	1716	911	53.1
East Germany	0.50	0.04	*35*	0.00	*94*	0.08	*8*	1362	1020	74.9
West Germany	0.62	-0.10	*25*	-0.12	*36*	-0.05	*35*	2042	734	35.9
Panel C										
Effect heterogeneity (living close to sports ground)										
Boys	0.57	-0.02	*63*	-0.03	*55*	0.02	*69*	1725	1499	86.9
Girls	0.51	0.00	*93*	-0.01	*84*	0.05	*32*	1679	1450	86.4
3-6 years old	0.41	0.08	*10*	0.09	*10*	0.03	*60*	1688	1470	87.1
7-10 years old	0.52	0.02	*79*	0.01	*91*	0.05	*60*	1716	603	35.1
East Germany	0.43	0.09	*4*	0.11	*4*	0.05	*34*	1362	1328	97.5
West Germany	0.58	-0.06	*20*	-0.06	*28*	-0.09	*8*	2042	1616	79.1

Note: Note: p-values obtained from 999 bootstrap replications.

Figure A.1: Nonparametric estimates of the relation of the distance by road to different sports facilities and sports participation



Note: No further covariates included. Nonparametric kernel regression based on the Epanechnikov kernel with a bandwidth of 0.6 is used for both figures.

Appendix B: Further details on the matching estimator used

This appendix describes the baseline matching protocol used for the matching estimator.

Table B.1: A matching protocol for the estimation of a counterfactual outcome and the effects (ATET)

Step A-1	Choose one observation in the subsample defined by *d*=1 and delete it from that pool.
Step B-1	Find an observation in the subsample defined by *d*=0 that is as close as possible to the one chosen in step A-1) in terms of $p(x), \tilde{x}$. 'Closeness' is based on the Mahalanobis distance.
Step C-1	Repeat A-1) and B-1) until no observation with *d*=1 is left.
Step D-1	Compute the maximum distance (*dist*) obtained for any comparison between a member of the reference distribution and matched comparison observations.
Step A-2	Repeat A-1).
Step B-2	Repeat B-1). If possible, find other observations in the subsample of *d*=0 that are at least as close as $R * dist$ to the one chosen in step A-2). Do not remove these observations, so that they can be used again. Compute weights for all chosen comparisons observations that are proportional to their distance. Normalise the weights such that they add to one.
Step C-2	Repeat A-2) and B-2) until no participant in *d*=1 is left.
Step D-2	D-2) For any potential comparison observation, add the weights obtained in A-2) and B-2).
Step E	Using the weights $w(x_i)$ obtained in D-2), run a weighted linear regression of the outcome variable on the variables used to define the distance (and an intercept).
Step F-1	Predict the potential outcome $y^0(x_i)$ of every observation using the coefficients of this regression: $\hat{y}^0(x_i)$.
Step F-2	Estimate the bias of the matching estimator for $E(Y^0 \mid D=1)$ as: $\sum_{i=1}^{N} \frac{d_i \hat{y}^0(x_i)}{N_1} - \frac{(1-d_i) w_i \hat{y}^0(x_i)}{N_0}$.
Step G	Using the weights obtained by weighted matching in D-2), compute a weighted mean of the outcome variables in *d=0*. Subtract the bias from this estimate to get $\widehat{E(Y^0 \mid D=1)}$.

Note: *R* is set to 90 percentage points.

The parameter used to define the radius for the distance-weighted radius matching (*R*) is set to 90percentage points. This value refers to the distance of the worst match in a one-to-one matching and is defined in terms of the propensity score. Furthermore, whenever single weight has a share larger than 6% compared to the overall sum of weights, the corresponding observation was removed (see Huber, Lechner, and Wunsch, 2010) for details.

When estimating the effect for the non-treated the same protocol is used but the role of treated and non-treated is reversed. Finally, the ATE is the weighted sum of the ATET and the ATENT where the weights are the share of the treated and the share of the non-treated, respectively.

Appendix C: Selection of estimation sample

Table C.1: Sample selection

Sample	No. of observations	Comment
Full sample	17,641	Full KiGGS dataset including children aged 0 to 17 years
Age group 3 - 10 years	8,023	Only observations of children aged 3 to 10 years are used as information on sports participation in- and outside clubs is only available for them.
Distance to facilities available	7,473	Number of observations where the distance could be estimated.
Sports participation available	6,998	Number of observations with non-missing answers on the questions regarding sports participation.
Final sample	6,998	