Individual labor market effects of local public expenditures on sports ☆

Tim Pawlowski [a],[*], Carina Steckenleiter [b], Tim Wallrafen [a], Michael Lechner [b],[#]

[a] *University of Tübingen, Institute of Sports Science, Wilhelmstraße 124, DE 72074 Tuebingen, Germany*
[b] *University of St. Gallen, Swiss Institute for Empirical Economic Research (SEW-HSG), Varnbüelstrasse 14, CH 9000 St. Gallen, Switzerland*

ARTICLE INFO

ABSTRACT

By merging administrative data on public finances of all municipalities in Germany with individual data from the German Socio-Economic Panel, we explore whether local public expenditures on sports facilities influences individual labor market outcomes. Our identification strategy follows a selection-on-observables approach and exploits the panel structure of the data covering 12 years between 2001 and 2012. The results of our matching estimations suggest that *both* women and men exposed to high annual expenditure levels (i.e. €31–€85 per capita) over 5 years, obtain approximately €150 of additional household net income on average. However, this income effect is captured by earning gains for men rather than for women living in the household. Additional analysis suggests, that these gender differences, which can also be observed in terms of working time, hourly wage and employment status, appear plausible since women in the age cohort under consideration are less likely to engage in sports in general, and in any of the publicly funded sports facilities in particular. Moreover, improved well-being and health are possible mechanisms through which the positive labor market effects for men may unfold.

1. Introduction

Although the literature analyzing the effects associated with public expenditures on health (e.g., Cremieux et al., 1999), education (e.g., Jackson et al., 2016) and labor market programs (e.g., Ham et al., 2011) is rich, little is known about related effects of public expenditures in other areas such as sports. This is surprising, given that sports participation can improve health (e.g., Charness and Gneezy, 2009), educational success (e.g., Stevenson, 2010), as well as labor market outcomes (e.g., Lechner, 2009). Furthermore, policy makers commonly legitimate public expenditures on sports with the explicit objective of fostering such effects (e.g., Federal Ministry of Interior, 2010).

We use nationwide micro-level panel data for Germany to explore whether sports-related public expenditures (SPEs) at the municipality level may lead to labor market success at the individual level. In Germany the total amount of annual public expenditures on sports is estimated at approximately €10 billion. The largest shares (i.e. about €3-4 billion each) are spent on physical education (P.E.) teachers in schools as well as the construction and maintenance of sports infrastructure (Pawlowski and Breuer, 2012). We focus on the latter since

sports infrastructure forms the basis for sports activity across *all* age groups. Recently, it became an explicit policy target in order to reduce physical inactivity in society.[1] We focus on the lowest administrative level, that is, municipalities, because money spent by regional authorities on sports facilities is commonly channeled through local authorities.

The expenditure data come from the accounting records of more than 12,000 municipalities in Germany and are matched with meteorological data, information on the socioeconomic characteristics of the local population and extensive individual data from a representative panel of German households, the German Socio-Economic Panel (SOEP) (2014). To consider that individuals may not only benefit from the expenditures of their own municipalities but also neighboring ones, we construct a distance-weighted expenditure measure.

The rich data set allows us to follow a selection-on-observables approach and exploit the panel structure of our data to identify the effects of interest. In this regard, we control for regional (e.g., population size, population density, weather conditions and geographic information) and individual characteristics (e.g., educational level, family status and household composition) expected to jointly influence public expenditures and labor market outcomes. Moreover, by conditioning on pre-exposure sports participation, labor market performance and various health and social capital indicators, we are able to indirectly control for unobservable individual traits confounding the relation of interest.

Overall, our matching analysis reveals sizable effects of local sports facility expenditures on individual labor market outcomes. For *both* men and women, we observe an economically and statistically significant increase in average household monthly net income of approximately €150 when moving from medium to high levels of SPE.[2] These effects are captured by earnings gains for men rather than women living in the same household.[3] Such gender differences can also be observed in terms of working time, hourly wages and employment status. Additional analysis based on the SOEP-Innovationssample (SOEP-IS) (2019) suggests that such effects may plausibly occur.[4] Women in the age cohort under consideration are less likely than men to engage in sports *in general* and to use any of the publicly funded sports facilities *in particular*.

Intuitively, one might expect that labor market effects occur through the following channels: comparably larger SPE increase the quantity and/or quality of sport infrastructure inducing changes in sports participation patterns of residents (e.g., Humphreys and Ruseski, 2007). If these changes are sufficiently large, social capital might increase (e.g., Schüttoff et al., 2018) and health might improve (e.g., Warburton et al., 2006). In the end, as argued by Lechner (2009), positive labor market effects could unfold through increased productivity, improved social networking and/or signaling of good health and motivation. Such effects would reduce, if (some) individuals decide to reduce their working time in order to practice (more) sports (Becker, 1965). At the same time, multiplier effects might occur if SPE-induced earning gains by residents are

spent on local products and services. Next to such *individual* channels, other (rather) *institutional* channels might also be plausible. For example, well-equipped sports infrastructure might serve as a soft factor in attracting firms (e.g., Porter, 1998). Moreover, if local firms are contracted for sports facility-related renovation or maintenance work, one might observe some SPE-induced earning gains for individuals working in these firms.

In further analyses, we find some indirect evidence that improved well-being and health as well as an increase in social capital are possible mechanisms that determine how the labor market effects we found may unfold. As such, while recent evidence suggests that higher levels of SPE do not increase the probability to become active (Steckenleiter et al., 2019), our findings suggest that higher levels of SPE may increase activity levels of already active individuals. Since, however, individual gains in social capital and health do not fully explain the effects we find, some other (*institutional*) channels might also be relevant, though to a lesser extent.

The structure of the paper is as follows. We start with discussing the related literature in Section 2. Section 3 presents the institutional setting and the sports-related expenditure data. Section 4 outlines the econometric specification, including details about the sampling, identification and estimation strategies. Section 5 discusses the findings. Section 6 summarizes the findings and offers some conclusions. Appendix A provides more information on the data collection. Appendix B provides a description of variables and additional estimation results, including information about the operational characteristics of the estimators. Appendix C contains some robustness checks. Appendix D reports results on possible mechanisms that determine how the effects may unfold.

2. Related literature

The literature analysing the correlates, determinants and effects of physical activity, sports participation and exercise (i.e. leisure time physical activity, LTPA, from now on) is vast and already subject to several reviews.[5] In this section, we focus on studies that have previously explored (parts of) our supposed channel, i.e. the link between public expenditures and/or the provision of sports facilities and LTPA as well as the effects of LTPA on labor market success.

Most of the only few studies which have explored the link between public expenditures and LTPA are based on highly aggregated data. For instance, Humphreys and Ruseski (2007) study the link between public expenditures in the 50 US-states and sports participation based on data provided by the Behavioral Risk Factor Surveillance Systems. They find that outdoor activities of the population are positively associated with public expenses on parks and recreation. Likewise, Dallmeyer et al. (2017) find a positive association between public expenditures on sports facilities and swimming pools in the 16 states ('*Länder*') and regular sports participation in Germany. In contrast, Kokolakakis et al. (2014) do not find any significant association between sports funding in the 325 local authorities and sports participation in England. Steckenleiter et al. (2019) provide the only study looking at the lowest administrative level, i.e. municipalities. They use the continuous nature of our SPE measure (see Section 3) in combination with individual sports participation information coming from the SOEP. Estimated dose-response functions reveal, that expenditures on sports facilities do not impact the probability to practice sports (we further discuss this finding in Section 5.5).

Most studies exploring the link between sports facility provision and LTPA have either employed a measure for *availability* (i.e. the number and/or density) or *proximity* (i.e. the physical distance) of sports facil-

---

[1] For instance, the World Health Organization recommends amongst others (2013, 34): "Creation and preservation of built and natural environments which support physical activity in schools, universities, workplaces, clinics and hospitals, and in the wider community, with a particular focus on providing infrastructure to support active transport, i.e. walking and cycling, active recreation and play, and participation in sports."

[2] Moving from medium to high levels of SPE translates into an average increase in annual spending of approximately €2.5 million for a medium-sized city with 75,000 inhabitants. This corresponds to a total investment with a present value of around €11 million over five years (which is the exposure period in our analysis).

[3] In general, this finding is reinforced by our results for the comparison between low and high SPE levels. However, the estimated effects are less precise due to some technical issues.

[4] The SOEP-IS was established in 2012 and covers more than 3,000 households. It included a module on physical activity in 2013, 2015 and 2017 (Richter and Schupp, 2015).

[5] See, for instance, Cabane and Lechner (2015), Downward et al. (2009), Downward et al. (2019) or Lechner (2015).
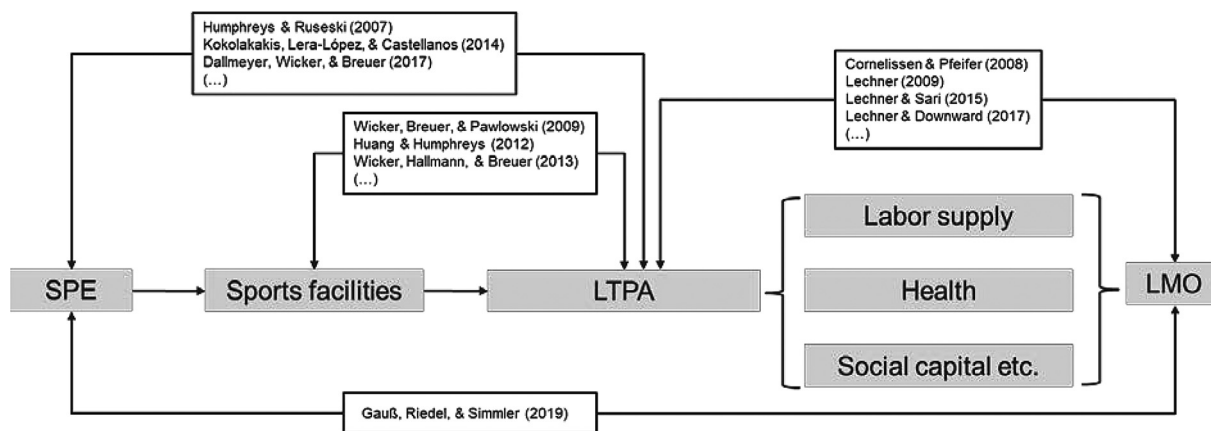
Fig. 1. Related literature exploring some underlying channels supposed for this study.
Notes: LMO: individual labor market outcomes. LTPA: leisure time physical activity. SPE: sports-related public expenditures.

ities in a given region.[6] For instance, Wicker et al. (2009) use micro- and macro-level survey data from Stuttgart, a major city in Southern Germany. Their findings suggest, that the availability of sports infrastructure is positively associated with sports activity in the different suburbs. In a follow-up study, Wicker et al. (2013) use geo-coded information about sports infrastructure in combination with survey data from Munich, another major city in Southern Germany. While they confirm a positive association for the number of swimming pools in the neighborhood, their findings for other types of sports facilities are mixed. Using data from the Behavioral Risk Factor Surveillance Systems, Huang and Humphreys (2012) find that greater access to sports facilities in a county increases sports participation in the US. As such, they further use this measure of sports facility provision for instrumenting sports participation in their happiness equation. Studies employing proximity measures have mainly focused on adolescents. For instance, Steinmayr et al. (2011) use data from the German Health Interview and Examination Survey for Children and Adolescents for estimating the effects of distance to the closest sports facility on sports activities inside and outside of sports clubs. They find that distance matters in smaller towns, particularly on the countryside.

Finally, several studies have already explored the link between LTPA and labor market outcomes. Since certain levels of LTPA are commonly associated with improved mental and physical health (Warburton et al., 2006) as well as an increase in social capital (Schüttoff et al., 2018), labor market effects may unfold, for instance, through increased productivity, signaling of good health or improved social networking (Lechner, 2009). Cornelissen and Pfeifer (2008) use data from the SOEP to explore the effects of sports participation during adolescence and as an adult on earnings in Germany. Overall, they report positive and significant effects. However, effect sizes vary by type and frequency of sports practiced as well as by gender. For instance, men involved in competitive sports during adolescence earn 6.4% (3.5%) more than men who were inactive (or not involved in competitive sports) dur-

ing adolescence. Likewise, men who practice sports at least weekly (at least monthly) earn 5% (2%) more than inactive men. While they also report significant gains in earnings from both competitive and non-competitive sports during adolescence for women (i.e. 5-6%), they do not find such effects for adult sports participation. This is in contrast to Lechner (2009), who exploits the panel structure of the SOEP and employs an identification and estimation strategy similar to ours (see Section 4). He reports an increase of annual gross earnings for both sports active men and women of about €1,200. Follow-up studies in other countries report even larger gains. For instance, Lechner and Sari (2015) use data from the Canadian National Population Health Survey and find a gain in earnings of 10-20% when moving from moderate to more intense activity levels. Likewise, Lechner and Downward (2017) exploit data of the Active People Survey commissioned by Sport England and report a gain in annual household income between £4,300-6,500 (£3,400-5,300) for men (women) dependent on the type of sports practiced.

Summing up, previous studies provide some correlational (and partly causal) evidence for the relevance of (at least some of) the channels supposed for our analysis. Gauss et al. (2019) provide the only study that directly investigates (amongst others) the link between local spending on sports facilities and labor market outcomes by using city-county expenditure data in combination with German linked employer-employee data. More precisely, they explore the effects of local public good and service provision on workers' wages in Germany. While their baseline results suggest that the stock of recreation and sports facilities does not have any significant effect on wages, further analysis suggests that a higher stock of recreation and sports facilities may even reduce wages of high-skilled workers. Importantly, however, their analysis is focused on urban areas which "in general provide a high level of infrastructure" (as acknowledged on p. 2). Moreover, they focus on short-term effects while findings by Lechner (2009) or Lechner and Sari (2015) suggest that changes in sports participation patterns translate into labor market effects after several years. As such, we provide the first study exploring the *long-term* effects of *local* public sports facility expenditures on *individual* labor market outcomes (see Fig. 1 for an overview of the related literature).

3. Sports-related public expenditures: data and measure

Sports in Germany are subsidized at all administrative levels, that is, by federal, regional and local authorities.[7] Regional and local authorities

---

[6] A systematic review of national sports strategies in 15 EU member states suggests, that the quality of sports infrastructure matters for LTPA, since providing quality sport facilities is frequently mentioned as a core objective in these documents (World Health Organization, 2009). Although the literature exploring the link between the quality of infrastructure and physical activity is generally scarce, there is some evidence suggesting that recreational activity is not only positively associated with access to recreational facilities but also objective measures of attractive features (Hoehner et al., 2005). Moreover, Brink et al. (2010) show, that renovated schoolyards increase physical activity at both an extensive and intensive margin, i.e. the number of children who are physically active increases, as well as their overall activity levels. More recent findings suggest, that the quality of school sports facilities may even influence physical activity patterns during adulthood (see Black et al., 2019).

[7] Note that Steckenleiter et al. (2019) use the same data analyzing the impact of SPE on the probability that an individual will engage in sports (see the discussion in Section 5.5). Therefore, Section 3 has certain overlaps with the data description in Steckenleiter et al. (2019). However, including this information

in particular aim to foster leisure sports participation (Pawlowski and Breuer, 2012). Because money from regional authorities is commonly channeled through local authorities, we focus our analysis on the lowest administrative level: district and urban municipalities/cities, including Hamburg, Bremen and Berlin which are independent from any of the other 13 states (*Länder*) in Germany.

In general, spending on sports is a voluntary self-government task (*freiwillige Selbstverwaltungsaufgabe*) in Germany, i.e. each municipality may independently decide about whether and how to finance sports. Revenues for such self-government tasks regularly come from (state) transfers and taxation, particularly property and business taxes as well as the municipalities' shares of income and value-added taxes.

Expenditure data from these more than 12,000 municipalities come from the Research Data Center (RDC) of the Federal Statistical Office and Statistical Offices of the Länder (2018a) and cover the financial years between 2001 and 2006. They include expenditures for the construction and maintenance of leisure sports facilities, sports airfields, sports stadiums, sports fields, tennis courts, toboggan and bobsled runs, sports schools, gyms, ski jumps, indoor or outdoor pools and Olympic sports facilities (Federal Statistical Office, 2016). Overall, there existed around 231,000 sport facilities in 2012, including facilities such as bowling centers, miniature golf areas, pool billiard rooms and (private) fitness centers (Federal Ministry of Economic Affairs and Energy, 2012). *Core* sport facilities (as defined by the Sports Minister Conference (2003)) amount to 136,000. Hereof, around 50% are (uncovered) sports fields such as football fields, outdoor tennis courts and swimming pools, another 25% are '*general*' indoor sport facilities, while around 10% (5%) are indoor tennis courts (indoor swimming pools). Comparing the available estimates for the number of *core* sport facilities at the national level between 2012, i.e. 136,000 (Federal Ministry of Economic Affairs and Energy, 2012), and 2000, i.e. 124,000 (Sports Minister Conference, 2003), suggests that the total number of sport facilities remains rather stable over time.[8] Even though cycling, jogging, or hiking belong to the most often practiced leisure type activities in Germany, representative survey evidence suggests that other sports – regularly practiced in such facilities – are also very popular.[9]

To work with the expenditure data, we converted all figures to 2004 Euros using the Consumer Price Index from the Federal Statistical Office (2017) and then cleaned, transformed and enriched the data in several steps. For example, since an increasing number of municipalities have transformed sports facilities (particularly swimming pools) into owner-operated community enterprises over time, we added financial data for all such sports-related enterprises from the annual balance sheets of public funds, institutions and enterprises in Germany (RDC of the Federal Statistical Office and Statistical Offices of the Länder, 2018b). Moreover, because some funded sports facilities (particularly soccer stadiums) are predominantly used by professional sports clubs only, we subtracted income from rental fees. The intuition behind this approach is that comparably higher construction and maintenance costs for sports facilities such as professional soccer stadiums are accompanied by comparably higher rental fees. In line with this, we reduced all figures in the 12 cities hosting any of the FIFA World Cup 2006 games by the amounts spent for the construction and renovation of stadiums for this event.[10]

Since individuals may not only benefit from the expenditures (and the resulting services and facilities) of their own municipalities but also neighboring ones (though to a lesser extent due to higher travel costs), we construct a distance-weighted expenditure measure instead of just taking the expenditure from the municipality of residence. Our measure considers, for an individual k living in municipality m, per capita expenditures of the municipality of residence D_m as well as the sports-related expenditures of all neighboring municipalities D_i within a certain radius around the municipality of residence:

$$D_k = \frac{x}{r} D_m + \left\{ \frac{(r-x)}{r} \left[\frac{\sum w'_i D_i}{\sum w'_i p_i} \right] \right\},$$

where x measures the distance between the geometric centers (calculated as polygon centroids) of the closest neighboring municipality and the municipality of residence, r measures the radius, p measures the population size of the neighboring municipalities, and $w'_i = (1/d_i)/(\sum 1/d_i)$ as further weight with d measuring the distance to the geometric center. In summary, this measure ensures that (i) the shares of expenditures for all municipalities under consideration for constructing each D_k add up to 100 percent and that the expenditures of the municipality of residence is weighted more (ii) the larger the municipality of residence is and (iii) the farther away the neighboring municipalities are.[11]

Fig. 2 shows the distribution of these distance-weighted per capita expenditures (five-year average, 2002–2006) across Germany after discretizing our measure by cutting off at the 33rd and 66th percentiles. Because some outliers are present, we kept observations with positive SPE values only up to the 99th percentile in this figure as well as in our estimations.[12] While we observe comparably larger amounts spent in Baden-Württemberg (located in southwest Germany), a significant within-state variation in SPE is generally evident across Germany. In line with Harris et al. (2001), this underscores the relevance of analyzing the lowest administrative level in our study.

4. Empirical strategy

We explore the effects of local SPEs from 2002 to 2006 on labor market outcomes in the period 2007–2012 by following and comparing individuals with similar characteristics in 2001 over time. We do so by merging our distance-weighted expenditure measure as defined in Section 3 with individual information from the SOEP (2014)[13] on a *yearly basis*. Merging on a yearly basis allows us to allocate 'true' expenditure values to individuals who relocated within the exposure period 2002–2006 before averaging the expenditure measure over five years. Averaging the expenditure measure seems reasonable to smooth out any sports facility 'investment shocks' that might occur over time.

To work with these panel data, we have to impose a couple of sample restrictions (see Section 4.1). Moreover, to address selection and endogeneity issues, we exploit the panel structure of our data and

---

is indispensable to fully shed light on the empirical strategies employed in both papers.

[8] This assumption is reinforced by the fact that both figures are quite comparable in size although both types of data collection methods were completely different, i.e. top-down in 2000 vs. bottom-up in 2012.

[9] For instance, 36% of the total population practice swimming (hereof, 37% practice at least once per week). Moreover, 15% of the total population practice gymnastics (hereof, 72% practice at least once per week) while another 13% of the total population plays soccer (hereof, 33% practice at least once per week) (Federal Ministry of Economic Affairs and Energy, 2019).

[10] For more information on the data cleaning, transforming and enriching process, see Appendix A.

[11] Further details on the construction of this measure are provided in Appendix A.

[12] By construction, our expenditure measure has negative values for municipalities exhibiting sports facility–related income (i.e., rental fees) higher than the corresponding (running) costs. Moreover, some municipalities spent considerably more than the others during the observation period. We decided to delete these municipalities from our analysis (denoted as gray areas in Fig. 2). Note that while Fig. 2 is based on all municipalities in Germany, the estimation sample consists of only the subsample of municipalities matching any of the SOEP observations.

[13] The SOEP is a longitudinal panel survey containing a representative sample of German households since 1984 (Goebel et al., 2018). Data are merged by using the official residence municipality key code (Amtlicher Gemeindeschlüssel – AGS). Because SOEP data, including the official residence municipality key code, are subject to severe restriction policies, data merging and analysis took place in the RDC of the German Economic Institute (Deutsches Institut für Wirtschaftsforschung – DIW) in Berlin.
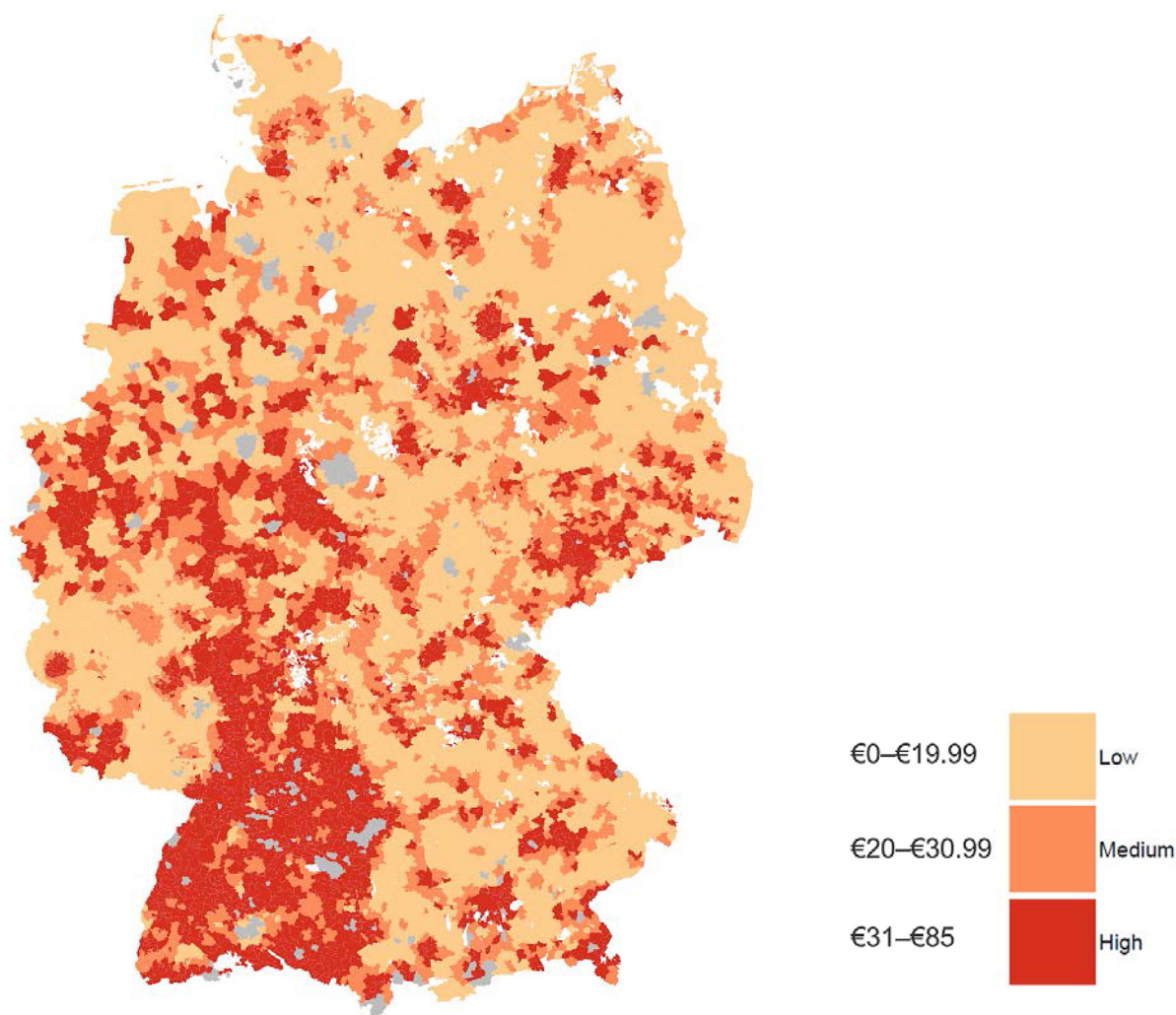
Fig. 2. Distance-weighted sports facility per capita expenditures in Germany.
Notes: Mean (2002–2006) net values (i.e., based on expenditure grouping number 999 minus income grouping number 399) in 2004 euros. Gray areas display municipalities with corresponding expenditure levels outside this range. White areas are uninhabited areas such as forests, lakes etc. Sources: RDC of the Federal Statistical Office and Statistical Offices of the Länder (2018a, 2018b), own calculations. The shape file for this map is provided by © GeoBasis-DE / BKG 2017.

apply a selection-on-observables approach by controlling for the relevant variables jointly influencing local sports expenditures and our outcome variables. These confounders include several *regional characteristics* (e.g., population size and density, weather conditions and geographic information) as well as *individual characteristics* (e.g., sociodemographics; information on sports participation behavior; various education, health, social capital and labor market indicators) taken from the pre-exposure year 2001 (see Section 4.2). To estimate the effects of interest, we employ radius matching with bias adjustment (see Section 4.3).

4.1. Sampling

Similar to Lechner (2009) and Lechner and Sari (2015), we are predominantly interested in *long-run* labor market outcomes. A long observation window, however, comes along with two problems: greater panel attrition and the increasing probability of retirement for older individuals.[14] Therefore, we decided to restrict our analysis to a six-year (post-

exposure) observation window between 2007 and 2012. As such, our panel data cover 12 years: the pre-exposure year 2001, an exposure period between 2002 and 2006 and a post-exposure period between 2007 and 2012, during which we estimate the labor market effects.

To work with this panel, we *first* restricted our sample to individuals who had a valid interview in 2001 and the first outcome year under consideration (i.e., 2007). At this stage, we also excluded individuals with missing values in lagged expenditure, lagged outcomes or any confounding variable used in our models (see Section 4.2). *Second*, we restricted our sample to all individuals aged 20 years and older in the pre-exposure year and also excluded individuals not yet graduated in 2001. *Third*, we imposed an upper age limit of 52 years in 2001 because these individuals would be aged 63 years in 2012 and as such still below the average age of retirement (German Statutory Pension Insurance Scheme, 2017). Moreover, we excluded individuals who were still in the education system from 2007 onward (e.g., trainees, students, civilian servants, conscripts). *Finally*, to avoid the results being driven by outliers in expenditures, we restricted the sample to observations up to the 99th percentile and only kept observations with positive values for the expenditure variable. After imposing all restrictions, the data set consists of 3,427 (3,071) women (men) living in 1,414 (1,265) different municipalities (see Table 1).

---

[14] Although we are able to test for any panel attrition in our study (see Section 5.3), smaller samples generally suffer from efficiency loss.

Table 1
Estimation samples and total population in Germany per state in 2007.

	Estimation samples				Total population
	Men		Women		
	N	%	N	%	%
Baden-Wuerttemberg (BW)	351	11.4	372	10.9	13.0
Bavaria (BY)	423	13.8	476	13.9	15.3
Berlin (BE)	104	3.4	122	3.6	4.0
Brandenburg (BB)	139	4.5	150	4.4	3.1
Bremen (HB)	23	0.8	27	0.8	0.8
Hamburg (HH)	33	1.1	50	1.5	2.1
Hesse (HE)	210	6.8	249	7.3	7.4
Mecklenburg-Western Pomerania (MV)	59	1.9	79	2.3	2.1
Lower Saxony (NI)	266	8.7	287	8.4	9.7
North Rhine-Westphalia (NW)	644	21.0	736	21.5	21.9
Rhineland-Palatinate (RP)	155	5.1	170	5.00	5.0
Saarland (SL)	39	1.3	40	1.2	1.3
Saxony (SN)	255	8.3	278	8.1	5.1
Saxony-Anhalt (ST)	134	4.4	158	4.6	3.0
Schleswig-Holstein (SH)	78	2.5	83	2.4	3.5
Thuringia (TH)	158	5.1	150	4.4	2.8
Total	3,071	100	3,427	100	100

Notes: Sources: Federal Statistical Office (2019), RDC of the Federal Statistical Office and Statistical Offices of the Länder (2018a, 2018b) and SOEP; own calculations.

4.2. Identification

When exploring whether local public expenditures on sports facilities influences individual labor market outcomes empirically, selection and endogeneity issues are obvious concerns. *First*, we expect *confounding issues* at both the aggregate (municipality) and individual level. For example, local authorities may use observed sports activity levels to determine SPE levels. Likewise, local labor market characteristics or weather conditions may influence both SPE levels and individual labor market outcomes. Moreover, individuals may deliberately make location choices based on the availability and quality of sports infrastructure (a behavior that can be described as a kind of Tiebout (1956) sorting effect).[15] *Second*, we anticipate a *reverse causality* issue since differences in labor market performances of residents might alter (income tax related) revenues of municipalities. This in turn might influence local SPE levels. *Third*, comparably richer municipalities might just spend more money on *all* voluntary self-government tasks including sports. In this regard, our analysis might suffer from a *simultaneity issue* since expenditures in other areas than sports might also influence individual labor market outcomes. Our approach, to consider a multi-year exposure period in order to smooth out any sports facility 'investment shocks' over time, might further aggravate the latter two concerns.

Since we could not use arguably exogenous variations in SPE as instrument, we exploited the panel structure of our data and applied a *selection-on-observables* approach to address these concerns. The key idea of this approach is that after conditioning on *observed* confounding variables, selection into the level of SPE is random (Imbens, 2004). To ensure that the conditional independence assumption (CIA) holds, we must control for all variables jointly influencing SPE levels and individual labor market outcomes. These confounders include various *regional* and *individual characteristics* measured in the pre-exposure year 2001 (if not stated otherwise).

The *regional characteristics* were selected to control for any confounding influences at the *aggregate* (municipality) level. (i) The first variable we consider is population size as a relevant predictor for both the availability and conditions of local (sports) infrastructure[16] as well as the labor market conditions and prospects of individuals. To better consider the latter aspect, we use population size in the corresponding BIK region instead of population size in the municipality only.[17] Similar arguments hold with regard to population density (i.e., rural vs. urban areas). Since population density is not directly considered in the BIK regions, we also control for (ii) population per square kilometer and (iii) its interaction with the different BIK regions.

To directly control for the economic conditions as well as the level of education in the municipalities, we include (iv) the share of unemployed people in the working age population (v) and the share of graduates with higher education entrance qualification of total graduates.[18] While it is common practice in the public spending literature to control for sociocultural conditions in the municipalities in general (e.g., Busemeyer, 2007), controlling for the composition of the population with migration background is particularly relevant in our setting given the popular policy claim that sports can serve as a tool for integration (e.g., German Olympic Sports Confederation, 2019). Although we are not able to implement any fine-grained measure here, we control for (vi) the share of foreigners in a municipality.

We also expect weather conditions to influence the investment and maintenance costs for (sports) infrastructure in a municipality as well as the labor market outcomes of individuals. Although the literature has extensively documented the effects of temperature, precipitation and windstorms on labor productivity and other economic outcomes (for a review, see Dell et al., 2014), to the best of our knowledge, no study has addressed these variables' possible effects on infrastructure investments and maintenance costs before. However, it seems plausible to assume that practicing outdoor sports is less popular in rainy areas and that providing opportunities for outdoor sports such as public parks or green

---

[15] Unfortunately, there is no comparable administrative data on local private sports facility expenditures or infrastructure available. Simple correlations between the number of fitness centers per 1,000 inhabitants in 2006 (see DSSV, 2007) and the average SPE radius (pure) measure in the 16 states ('Länder') are low and not significant, i.e. .32 (.39), suggesting that there are no interaction effects between public and private sports infrastructure.

[16] We expect that, in general, larger municipalities have better opportunities to build and maintain such infrastructure because of economies of scale.

[17] A BIK region is made up of all municipalities for which at least 7 percent of its employed inhabitants (as identified by social insurance contributions) commute to the same core city (BIK Aschpurwis & Behrens, 2001). We take BIK regions with 500,000 and more inhabitants as a reference category and differentiate between three additional categories in our models: BIK regions with 0–49,999, 50,000–99,999 and 100,000–499,999 inhabitants.

[18] To avoid picking up any year-specific fluctuation, we take the six-year average (2001–2006) for both measures.

areas is less cost-intensive than providing and maintaining indoor sports facilities. Therefore, we also control for (vii) the average sum of precipitation and (viii) the average sunshine duration per year.[19]

Finally, given the significant regional variation in sports facility spending – with comparably higher expenditures in southern Germany (see Fig. 2), which is also a comparably more prosperous region and has better labor market conditions and prospects for individuals – we control for (ix) the region where the municipality is located: south (including the states Baden-Wuerttemberg and Bavaria), north (including the states Lower-Saxony and Schleswig-Holstein as well as the city states Bremen and Hamburg), west (including the states of Hessen, Rhineland-Palatine and Saarland) and east (including the states of Brandenburg, Mecklenburg-West Pomerania, Saxony, Saxony-Anhalt and Thuringia as well as the city-state Berlin).

Whereas we use regional characteristics to control for any confounding influences at the *aggregate* (municipality) level, we condition on *individual characteristics* (as measured in the SOEP) that are expected to influence both the labor market performance and the level of SPE at the *disaggregate* (individual) level. In this regard, we do not expect any individuals (other than the decision makers in the local government) to directly influence the level of SPE. However, (sports active) individuals may deliberately make location choices based on the availability and quality of local sports infrastructure.[20]

To systematically address this potential confounding issue, we control for those characteristics that previous research has shown to jointly influence individuals' sports participation behavior and labor market performance (for an overview, see Lechner, 2009), such as (i) being married, (ii) being divorced, (iii) having German citizenship, (iv) the number of children in the household and (v) age. In addition, we include two sets of educational measures. The first set considers the (vi) highest school degree (measured as lower secondary school or no school degree, intermediate secondary school degree and upper secondary school degree or other graduation diploma). The second set considers (vii) individuals' level of (higher) education/training (measured as still pursuing education, no vocational degree, vocational degree or university degree).

To directly control for differences in sports participation behavior in the pre-exposure period, we also include (viii) whether an individual participated at least monthly in sports, as indicated in the SOEP. Moreover, we consider differences in labor market performance in the pre-exposure period directly by including (ix) monthly household net income, (x) gross earnings, (xi) working time per week, (xii) hourly wage and (xiii) full-time employment.[21]

Finally, we also control for various health and social capital indicators from the pre-exposure period because they could influence sports participation behavior and labor market performance in the subsequent years. To this end, we control for (xiv) satisfaction with health, (xv) satisfaction with life, (xvi) the number of doctor visits per year and (xvii) the physical component (PCS) and (xviii) mental component summary scales (MCS).[22] The three available social capital indicators are dummies measuring interpersonal networks as (xix) whether an individual helped out friends, neighbors or relatives at least monthly and social engagement as (xx) frequency of volunteer work in clubs or social services (at least monthly) and (xxi) general involvement in a citizens' group, political party and/or local government.

While this rich set of regional and individual characteristics allows us to control for the most relevant *observable* confounders, we are unable to control for genetics or any psychological, cognitive or emotional factor that might influence sports participation behavior, location choice and individual labor market performance *directly*. However, we argue that conditioning on pre-exposure sports participation, labor market performance and the various health and social capital indicators enables us to *indirectly* control for such *unobservable* confounders at least to some extent. In this regard, the available measures 'control' for some sort of individual fixed effects as Lechner (2009) argues.

A natural question that arises at this stage is whether conditioning on these confounding variables is also helpful to control for the reverse causality and simultaneity issues mentioned at the beginning of this section. By employing simple OLS regressions we find, that the available (lagged) regional characteristics are able to predict between 81% and 87% of the variance in per capita income tax between municipalities in later years (see Table B.2 in Appendix B).[23] Therefore, we can provide some suggestive evidence that conditioning on these variables in the pre-exposure period also helps to mitigate the reverse causality issue mentioned before.

Unfortunately, we are not able to implement any comparably reliable empirical tests with regard to the simultaneity issue here, since *all* expenditures in *all* years are correlated by definition as long as the municipalities are budget constrained (which they obviously are). Nevertheless, the regression exercise mentioned before provides some suggestive evidence that conditioning on lagged regional characteristics enables us to control for (labor market related) revenues of the municipalities during the exposure period. As such, we are able to indirectly condition on the *general* spending capacities of the municipalities. Moreover, we argue, that substituting SPE with expenditure categories that do *not* alter labor market performance in society is per se not problematic while substituting SPE with expenditure categories that alter labor market performance would possibly just downward bias our results. As such, we would end up with rather conservative figures about the overall effects of SPE.

Finally, one might be concerned that municipalities alter local business tax in order to cover (additional) expenditures on sport facilities *or* spend more on sport facilities as a reaction to tax changes.[24] Both issues are not testable with the available data but might threaten our identification strategy, since workers were found to bear about one-half of the total tax burden (Fuest et al., 2018). The authors of this

---

[19] Because weather conditions are subject to changes in both the short and the long run, we decided to consider long-term averages covering the period 1981–2010.

[20] While we do not know anything about the extent to which leisure-time sports facilities might stimulate location choices in practice, some evidence shows that income and family status might increase the probability of living closer to professional sports facilities (e.g., Ahlfeldt and Kavetsos, 2014). Moreover, several studies investigating the socioeconomic determinants of location choice in general show some evidence for the United States that comparably less populated areas have relatively higher marriage rates (Lichter et al., 1991) and that the probability of living in suburban areas increases with increasing age (Alba and Logan, 1991).

[21] We converted all monetary measures to 2004 Euros using the Consumer Price Index from the Federal Statistical Office (2017) and controlled for hourly wage in our estimations to allow for more flexibility. Since hourly wage is not readily available in the SOEP, we computed it by dividing monthly gross earnings by weekly hours ($\times$ 4.3).

[22] The PCS and MCS are superordinate scales available since 2002 on a biannual basis. Both scales had been derived by Andersen et al. (2007) from an internationally applied inventory of health measures (the so-called SF-12v2 indicators) using explorative factor analysis. Because data from these scales are not available for 2001, we took both measures from the 2002 SOEP survey.

[23] The available data allows us to test this relation for 11 consecutive years, i.e. 2001-2011. While we used population size in the corresponding BIK region in our main models (see Section 5), we could only use population size in the municipality of residence in these auxiliary regressions since information on BIK regions is only available at the DIW in Berlin. We argue, however, that BIK regions by construction better pick up the labor market conditions and prospects of individuals in the municipalities. As such, our figures on the variance explanatory power of the different models are likely to be lower bound estimates.

[24] Local business tax forms the most important revenue stream for German municipalities. The corresponding tax base is operating profits. The tax rate consists of two components: the basic tax rate (set by the federal government) as well as a local scaling factor (set by the local government/municipality). The municipality council votes each year on the next year's local scaling factor, which may remain unchanged, decrease or increase (Fuest et al., 2018).
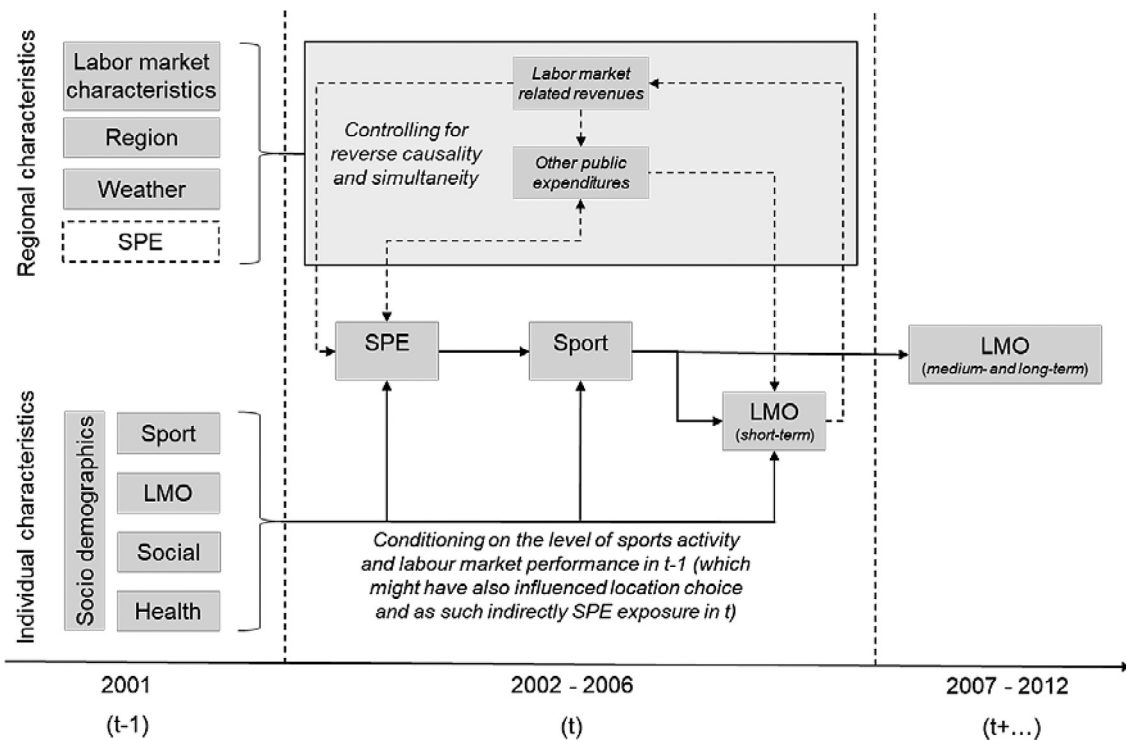
Fig. 3. Empirical design and the identifying assumptions.
Notes: Health: individual health indicators. LMO: individual labor market outcomes. Region: regional dummies. Social: individual social capital indicators. SPE: sports-related public expenditures at the municipality level. Sport: individual sports participation behavior. Weather: regional precipitation and sunshine hours. Detailed variable descriptions are provided in Table B.1 in Appendix B.

paper, however, find flat pre-trends when looking at municipality revenues and spending, suggesting that local tax rates do at least not respond to investment / expenditure "shocks". This is also in line with Foremny and Riedel (2014) who provide evidence that tax changes are triggered rather by political factors. Moreover, since not only cross-sectional variation in local tax rates but also variation over time reveals some regional clustering, this issue should be picked-up (at least to some extent) by our control variables.

Our empirical design and the underlying identifying assumptions are summarized in Fig. 3.

4.3. Estimation

We opt for a matching type estimator because it is more flexible and more robust with respect to the statistical assumptions imposed (Imbens, 2004) than fully parametric models. Similar to Lechner and Sari (2015), who estimate the labor market effects for three sports activity levels (active versus moderately active versus inactive), we discretize our SPE measure as defined in Section 3 in three groups: low (€0–€19.99), medium (€20–€30.99) and high (€31–€85), using the 33rd and 66th percentiles as cutoffs. Doing so allows us to compute and compare overall three average treatment effects (ATEs): the labor market effects of moving from low to medium, from low to high and from medium to high SPE levels.

Because we assume that CIA holds across all strata, the sample reduction results of Lechner (2001) apply; that is, for estimating any one of these three effects, participants in the irrelevant treatment state are deleted for the purpose of this particular estimation. However, doing so requires that observations *on support* for a certain group are the same across estimations. In other words, when estimating the effects of low versus high and medium versus high levels of SPE, the same observations must be considered in the treatment group (high SPE level). To ensure that this is the case, we implemented an iterative approach; that is, we deleted any observations off support after each estimation and re-

estimated all models (including propensity scores) until all remaining observations were on support.[25] Finally, all effects for women and men are estimated separately, because there are considerable gender-specific differences in terms of both sports activity patterns (e.g., Eccles and Harold, 1991) and labor market participation (e.g., Fitzenberger et al., 2004) .[26]

All effects are estimated by radius matching with bias adjustment, as Lechner et al. (2011) suggest. A simulation study (Huber et al., 2013) shows that this matching estimator performs particularly well.[27] We based inference on bootstrapping the sample 99 times.[28]

---

[25] A common support statistic of our estimated models is available in Table B.4 in Appendix B.

[26] Note that using the 33rd and 66th percentile cutoffs does not necessarily mean that we end up with an equal number of observations in each group for each model, because the SPE measure is discretized for the full sample (before splitting it into the men/women strata) and we consider only observations on support in the models.

[27] For computations, we used STATA and the corresponding command radiusmatch as developed by Huber et al. (2015). Our specification mainly follows the default options: a linear bias correction, a multiplier of 300 and a quantile of 90, indicating that the radius is equal to 300 percent of the maximum distance in pair matching and that we used the 90th quantile of the distances in pair matching. All analyses needed to be conducted in the RDC of the DIW in Berlin, because regional information in the SOEP is subject to severe data restriction policies. Unfortunately, the infrastructure at the DIW did not allow us to employ more advanced machine learning estimators.

[28] It might be questioned whether the number of bootstrap replications is sufficiently large. While we were not able to compute a larger number of bootstrap replications for all the different models estimated, we compared the results of one stratum (men, 2008) after bootstrapping 99 and 499 times. Overall, the results did not change much. This is in line with some more general evidence suggesting that reliable inference can be obtained with very small numbers (there: 49) of bootstrap replications (see Bodory et al., 2018). Therefore, we are not concerned about any serious problem in this regard.

Table 2
Mean values and marginal effects of the selection models (selected variables).

	Men 20–52 years in 2001						Women 20–52 years in 2001					
	Mean in subsample			Marginal effects			Mean in subsample			Marginal effects		
Characteristics (1)	High (2)	Med (3)	Low (4)	H–L (5)	H–M (6)	M–L (7)	High (8)	Med (9)	Low (10)	H–L (11)	H–M (12)	M–L (13)
Population size and density												
Population size ≥ 500k	.44	.38	.32	R	R	R	.46	.41	.32	R	R	R
Population size 100k–499k (L)	.18	.19	.12	-.20**	-.53***	.15	.17	.18	.11	-.04	-.27***	.05
Population size 50k–99k (M)	.02	.04	.07	-.09	-.41	.13	.03	.04	.08	-.002	-.19	.02
Population size 0–50k (S)	.35	.38	.49	-.11**	-.42***	.17**	.34	.37	.49	-.11*	-.37***	.10*
Population density	1.8	1.7	1.2	.02	-.14***	.10***	1.8	1.8	1.3	.02	-.10***	.06***
Population density × L	.38	.35	.17	.16***	.22***	.03	.32	.34	.17	.08*	.10**	.07**
Population density × M	.04	.07	.10	-.05	.09	-.06	.04	.07	.11	-.09*	-.008	-.03
Population density × S	.29	.22	.17	.08***	.18***	-.02	.29	.21	.20	.07**	.17***	-.008
Population characteristics												
Share of unemployed [1]	7.1	8.9	9.2	-.02***	.02	-.03***	7.2	8.9	9.4	-.02***	.01***	-.03***
Share of graduates with HEEQ [1]	24	25	23	.003	-.0003	.005**	24	25	23	.0009	-.002	.002
Share of foreigners	1.9	1.5	1.2	.03	.03	.005	1.9	1.5	1.2	.06**	.02	.04**
Weather conditions												
Average precipitation [2]	2.1	2.0	2.0	.33***	.28***	.06	2.1	2.0	2.0	.29***	.25***	.02
Average sunshine duration [2]	.47	.46	.45	.98***	.12	.80**	.47	.47	.46	.31	-.51	.34
Region												
North-Rhine Westphalia	.22	.20	.21	R	R	R	.25	.20	.20	R	R	R
Southern Germany	.40	.18	.17	.10**	.20***	-.05	.38	.19	.16	.12***	.22***	-.02
Northern Germany	.04	.14	.20	-.20***	-.27***	-.03	.03	.13	.21	-.29***	-.31***	-.09***
Western Germany	.19	.12	.08	.20***	.14***	.13**	.19	.12	.09	.16***	.15***	.07
Eastern Germany	.14	.35	.34	.13**	-.17**	.19***	.15	.36	.34	.17***	-.16**	.27***
Labor market indicators												
Household net income in EUR	2,733	2,577	2,592	.000009	.00001	-.00001	2,640	2,568	2,505	.00002*	.000005	.00001
Gross earnings in EUR	2,731	2,424	2,271	.000007	.00002	.00001	1,165	1,165	1,086	-.00002	-.00002	.000006
Working time per week	41	39	38	.0008	.001	-.003	23	23	23	.0006	.002**	-.001
Hourly wage in EUR	14	13	12	-.0003	-.004	-.0005	8	8	8	.003	-.001	.003
Employed full time	.88	.84	.81	.02	-.009	.04	.39	.39	.38	.019	.006	-.0008
	All other individual characteristics as described in Section 4.2 are included in the models and are displayed in Table B.3 in Appendix B.											
# of observations, pseudo-R²	985	1,051	1,009	22	13	7	1,095	1,150	1,135	21	12	8

Notes: All figures are based on the 2007 sample. Because we have an unbalanced panel with a decreasing number of observations due to panel attrition, we needed to re-estimate all probit models for each year to calculate the corresponding ATEs (results are available on request). The high (med) [low] sample consists of individuals living in a municipality with high (medium) [low] level of distance-weighted per capita SPEs as defined in Section 3. The dependent variables in the probit models are dummies measured as 1 for the relatively higher expenditure level (i.e., for the H–M-model high, for the H–L -model high and for the M–L-model medium. Independent variables are measured prior to the dependent variable if not stated otherwise. Significance levels are as follows: * ≡ $p \leq 10\%$, ** ≡ $p \leq 5\%$, *** ≡ $p \leq$ 1%. R denotes the reference category. A detailed description of all variables is provided in Table B.1 in Appendix B. Table B.3 in Appendix B reports the results for all variables, including the individual characteristics not displayed here. Sources: RDC of the Federal Statistical Office and Statistical Offices of the Länder (2018a, 2018b) and SOEP; further data sources as discussed in detail in Sections 3 and 4 and Appendix A; own calculations.

[1] Measured between 2001 and 2006.
[2] Measured between 1981 and 2010.

5. Results

We first describe our findings for the selection process (Section 5.1). We then turn to discussing the labor market effects (Section 5.2) as well as the results for the robustness checks (Section 5.3), effect heterogeneity (Section 5.4) and possible mechanisms (Section 5.5).

5.1. Selection process

Table 2 provides an overview on conditional mean values and marginal effects of the regional and individual characteristics used in the selection models based on the main specification.[29]

---

[29] To calculate the ATEs, we needed to re-estimate all probit models for each year because we have an unbalanced panel with a decreasing number of observations due to panel attrition. Note, that we opted against a balanced panel approach to avoid further efficiency loss in the first years of the post exposure period. Although all individual characteristics as described in Section 4.2 are included in all models, Table 2 only reports the results for the individual labor market indicators, which are the focus of our study. In general, we observe only a few significant differences in the individual characteristics (all results are reported in Table B.1 in Appendix B).

The data show that several mean values of the variables differ significantly between the three samples, which underscores the relevance of controlling for them in the first stage. A closer examination of the six models reveals that, in particular, the high versus low and high versus medium samples differ significantly. These differences are less evident when comparing the medium versus low samples, as reflected in the comparably lower pseudo-R² of the probit models for the medium versus low samples. Moreover, the observable patterns do not differ much between the male and female samples, suggesting that women and men with comparable characteristics are equally distributed in all subsamples.

With the exception of the medium versus low comparison, the data show that the larger the BIK region, the higher is the level of SPE. Moreover, the higher the population density (share of unemployed citizens in the working age population), the higher (lower) is the probability of observing medium rather than high or low SPE levels. In addition, the higher the share of foreigners in the municipality, the higher is the level of SPE, which might reflect a policy in line with the claim that sports can serve as a vehicle for integration. The coefficients are, however, only significant for women.

Furthermore, weather conditions and location influence the level of SPE as expected: Expenditures increase with increasing precipitation rates, and municipalities in the northern states spend comparably less

than North Rhine-Westphalia, whereas municipalities in the western and southern states spend comparably more. Moreover, the probability of observing low SPE levels in a municipality is lower in the eastern states than North Rhine-Westphalia. Finally, although individual labor market indicators suggest that living in municipalities with comparably higher SPE levels is correlated with comparably more success in the labor market, only 2 of 30 marginal effects are significant. This is in line with the observation that few significant differences in individual characteristics can be observed (see Table B.3 in Appendix B).

5.2. Labor market effects

We next turn to the estimated ATEs as described in Section 4.3. The following figures show the corresponding effects of moving from medium to high (Fig. 4), and from low to high SPE levels (Fig. 5) for monthly net household income, net earnings and gross earnings. Full sample estimates are derived from the weighted averages of the men and women sample estimates.

To see how the effects develop and whether (and to what extent) controlling for our confounding variables was successful, we also include the corresponding estimates for the pre-exposure (2001) and exposure periods (2002–2006) in each graph. Overall, the intuition is that balancing out the different confounding influences worked better the closer the estimated ATEs were to zero in 2001. A negative (positive) deviation from zero indicates a potential downward (upward) bias of our estimates. Accordingly, the balancing worked slightly better for men than for women with regard to household income and vice versa with regard to the earning measures.

Overall, while we do not observe any significant effects on income and earnings for moving from low to medium SPE levels (results are provided in Figure B.1 in Appendix B), we do find that men (women) have an *average*[30] gain in net household income of €153 (€157) when moving from medium to high SPE levels. The corresponding effects when moving from low to high SPE levels are less precise but of comparable size for men (i.e. on average €124).[31] Remarkably, further inspection of Figs. 4 and 5 suggests that these household income effects are captured by earning gains for men rather than women living in the household. While this is confirmed by our *mean* model estimates for the low versus high comparison, the corresponding estimate for the medium versus high comparison is not significant at the conventional level.[32]

Although the corresponding estimates are more volatile, we observe such gender differences also for other labor market indicators. While we do not find any significant effects in the post-exposure period on working time per week, hourly wage and the probability of being employed full time for women, we find some significant effects for men (see Figure B.2 in Appendix B).[33] For example, moving from medium to high SPE levels is accompanied by an average increase in working time of about half an hour in the post-exposure period. However, we only observe a significant estimate in 2009. Likewise, the corresponding effect in our *mean* model is not significant at the conventional level. Furthermore, we find a significant increase in hourly wage in the post-exposure period for all three models when moving to comparably higher SPE levels for some years. Considering the differences in the level of balancing in 2001 between the models, these effects seem to converge at approximately €1. However, according to our *mean* models, the average effect is only significant for the low versus high comparison (€1.2). Similarly, we observe an average increase of approximately 2–3 percent in the probability of being employed full time in the post-exposure period when moving to comparably higher SPE levels. However, these effects are only significant in 2009 for the low versus high and the medium versus high comparisons. Likewise, the *mean* model estimates are not significant.

In summary, we observe a significant (average) increase in household net income of approximately €150 in the post-exposure period when moving from medium to high SPE levels. These effects are captured by earning gains for men rather than women living in the household. This is confirmed by results for the low versus high comparison for men. Likewise, while our estimates for other labor market indicators are generally volatile, the gender differences we observe for earnings are broadly confirmed. Although we do not find any significant effects for women in the post-exposure period, our results suggest that (on average) men work about half an hour more and see a benefit of a €1 increase in wage as well as an approximately 2–3 percent increase in the probability of being employed full time when moving to comparably higher SPE levels.

5.3. Robustness checks

In order to evaluate the credibility of our selection-on-observables approach and, as such, the credibility of all estimates discussed before, we first consider the balancing statistics provided in Tables B.5 and B.7 in Appendix B. Overall, balancing out the different confounding influences worked quite well for the high versus medium (H–M) models. After condition on the propensity score, only few variables are left with a bias larger than 10%, while the overall mean biases considerably reduced, i.e. on average from 13.3% (11.5%) to around 4.4% (4.5%) for the men (women) samples. The mean biases also reduced considerably in the other two models. However, comparably more variables are left that still do not balance properly after conditioning on the propensity score. As a consequence, although several import variables (such as the corresponding lagged outcome variables) balance quite well in most cases, we have less confidence in the estimates of the high versus low (H–L) models and the medium versus low (M–L) models compared to the high versus medium (H–M) models.

To further test the credibility of our selection-on-observables approach, we implement placebo tests. As such, we re-estimate our mean models for two measures, i.e. the individual worries about (i) the protection of the environment or (ii) peacekeeping. Both measures are (the-

---

[30] If not stated otherwise 'average' refers to average post exposure effects (2007–2012). These effects were estimated using the mean of our outcome variables over the years during the post exposure period. We refer to these models as mean models in the following. ATEs from the mean models might deviate from just averaging the year-specific ATEs as observed in Figs. 4 and 5 since the mean models are based on the 2007-sample while the year specific models are based on the samples in the corresponding years. In other words, while year specific models are subject to panel attrition, the mean models are not since we just average over all available year specific observations for each panelist (in this regard the binary variable full-time employment measures 1 if the person was full-time employed in all years with observations). For instance, if a panelist drops out in 2010, she/he is not considered in ATE_{2010}, ATE_{2011} or ATE_{2012} of Figs. 4 and 5. However, she/he remains as observation in our mean models averaging over her/his outcomes just for 2007, 2008 and 2009 before estimation (see also Section 5.3).

[31] The zero-effect for women for the high versus low comparison can be attributed to less precise estimates and rather poor balancing as indicated by a significant ATE in 2001 (see Fig. 5). This is confirmed by some robustness checks for which the balancing worked slightly better and the ATE increases to around €100 (though still being insignificant at the conventional level due to large standard errors, see Section 5.3).

[32] Recall, that our mean models do not account for panel attrition. Since labour market effects unfold in the longer rather than the shorter run, our mean models more likely underestimate the long-term effects of interest. As such, these effects can be interpreted as kind of (averaged) lower bound estimates for the corresponding year specific models.

---

[33] As noted earlier, because hourly wage is not available in the SOEP, we approximated it by dividing monthly gross earnings by weekly hours (× 4.3). For non-workers, hourly wage measures zero. Theoretically, as acknowledged by a reviewer, the effect of SPE on working time is not clear given time constraints. If individuals have more opportunities to engage in sports (due to higher SPE), they might increase working time since they just feel better and vitalized. Likewise, they might reduce working time in order to practice (more) sports.
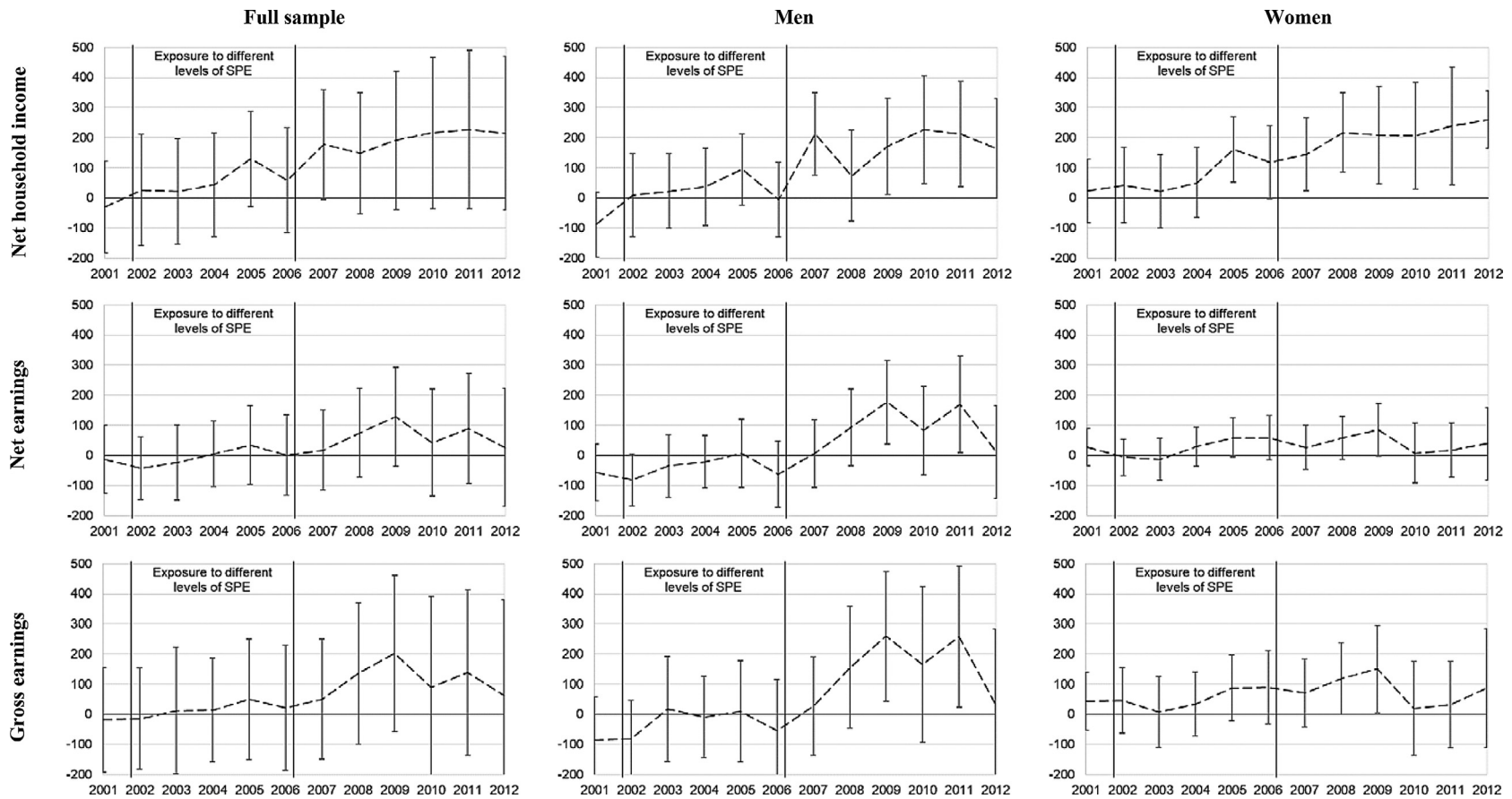
Fig. 4. Effects of high versus medium levels of SPE on household income and earnings.

Notes: Radius matching results for men and women as well as their weighted averages (full sample). This figure displays ATEs of high versus medium levels of distance-weighted per capita SPEs (as defined in Section 3) and the corresponding 95% confidence bands. All 2001 ATEs are calculated as mean differences between matched treated and untreated observations from the balancing tests of the 2002 estimations. We approximate the effects for net earnings in 2001 by multiplying the effects for gross earnings with the samples' average net-to-gross earnings ratio (i.e., 65.6% [63.5%] for men [women]). For common support statistics, see Table B.4 in Appendix B. Tables B.5 and B.7 in Appendix B provide an overview of the balancing statistics of all covariates for the estimated matching models from the post-exposure period (2007–2012). Sources: RDC of the Federal Statistical Office and Statistical Offices of the Länder (2018a; 2018b) and SOEP; further data sources as discussed in detail in Sections 3 and 4 and Appendix A; own calculations.

Fig. 5. Effects of high versus low levels of SPE on household income and earnings.

Notes: Radius matching results for men and women as well as their weighted averages (full sample). This figure displays ATEs of high versus low levels of distance-weighted per capita SPEs (as defined in Section 3) and the corresponding 95% confidence bands. All 2001 ATEs are calculated as mean differences between matched treated and untreated observations from the balancing tests of the 2002 estimations. We approximate the effects for net earnings in 2001 by multiplying the effects for gross earnings with the samples' average net-to-gross earnings ratio (i.e., 65.6% [63.5%] for men [women]). For common support statistics, see Table B.4 in Appendix B. Tables B.5 and B.7 in Appendix B provide an overview of the balancing statistics of all covariates for the estimated matching models from the post-exposure period (2007–2012). Sources: RDC of the Federal Statistical Office and Statistical Offices of the Länder (2018a; 2018b) and SOEP; further data sources as discussed in detail in Sections 3 and 4 and Appendix A; own calculations.
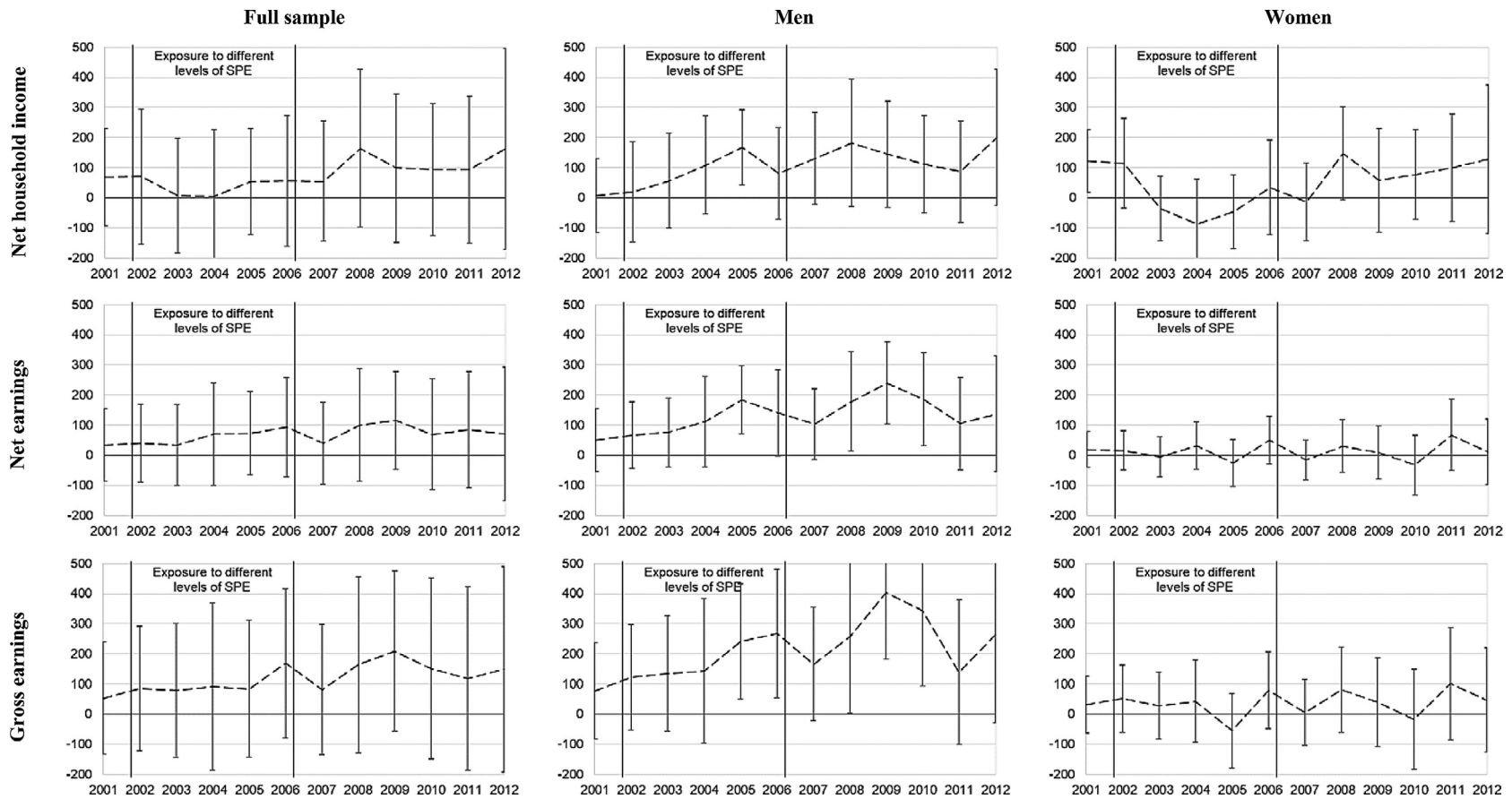
oretically) not influenced by SPE while they are at the same time not confounding the relationship between SPE and individual labor market outcomes. The intuition of testing such measures is as follows: if we would find any SPE-induced effects for these variables, our main findings could be subject to spurious correlation. Overall, we do not observe any significant differences between individuals in the pre-exposure period. Since we only observe few and marginally significant effects for the women sample in the post-exposure period, we argue that these tests do not cast any doubts on our main findings.[34]

We also test the sensitivity of our findings with regard to other sampling criteria, specifications and estimators. First, we test whether our results in the exposure periods are driven by individuals who were still in the education system between 2001 and 2006 because (for example) students are not randomly distributed across municipalities.[35] Although we expect the available regional and individual confounders to account for this issue, empirical testing seems advisable considering that the estimated ATEs in the pre-exposure period(s) already provide important information on the balancing and about how effects develop over time. Red lines in Figures C.1 and C.2 in Appendix C show the estimates for all samples *excluding* individuals who were still in the education system in 2001–2006 (e.g., trainees, students, civilian servants, conscripts; approximately 10–11% of the sample). Although some minor deviations in effect sizes emerged and balancing out the different confounding influences for the low versus high comparison worked better when individuals still pursuing education are *included*, the results do not indicate that our main findings are affected.

While this test has focused on the sensitivity of our estimated (non-)effects during the pre-exposure period, further robustness checks focus on the sensitivity of our effects in the post-exposure period. In order to facilitate understanding and comparability of the various tests employed, we present average ATEs from the corresponding *mean* models for all robustness checks as bars with confidence intervals in a single graph (see Figures C.3-C.5 in Appendix C).

In order to explore whether municipalities with big spending spikes during the exposure period might affect our findings, we re-estimate all models without truncation, i.e. including observations with positive SPE values above the 99th percentile (Test A). Likewise, in order to test whether our results are driven by individuals who move to places with higher levels of SPE during the exposure period, we re-estimate our

models excluding movers (Test B). While effect sizes increase – particularly for the high versus low comparison in the men strata – our main conclusions remain.

Test C additionally controls for SPE measured in the pre-exposure year 2001. In general, controlling for lagged SPE is certainly a comparably stricter condition that might add more credibility to our identification strategy. However, we avoided this approach as a main specification because it requires a sufficiently large variation between the 2001 and the (average of the) 2002–2006 SPE levels, both within subject and between subjects. Consequently, this approach emphasizes comparisons between treatment and controls close to the thresholds (i.e., between medium versus low and high versus medium SPE levels). Moreover, it reduces the probability of finding adequate matching partners for the extreme case, that is, the comparison between high and low SPE levels. Overall, these concerns are confirmed by the estimates of our *mean* models, particularly for the high versus low comparison (as expected).[36] At the same time, however, since earning gains for men remain substantially larger than those for women, this does not generally question our main conclusions.

Test D explores how changes in SPE relate to *changes* in our outcome measures by using the difference between $outcome_t$ and $outcome_{2001}$ as dependent variables. While effects sizes reduce for the high versus low comparison (particularly for men), effect sizes increase for the high versus medium comparison. Although average earning gains for women become marginally significant when moving from medium to high SPE levels, they remain half in size compared to the ATEs for men. As such, our interpretation that any household income effects are captured by earning gains for men rather than women living in the household is supported.

We also test whether pre-exposure *trends* in our outcome measures influence our findings by additionally controlling for the difference between $outcome_{2001}$ and $outcome_{2000}$ (Test E).[37] Overall, while effect sizes remain similar for the high versus low comparison, they increase for women and decrease for men for the high versus medium comparison. Importantly, however, further examination of the corresponding *mean* model fit for men reveals that conditioning on pre-exposure *trends* in our outcome measures considerably worsens the balancing statistics for several lagged outcome variables. Most notably, the mean biases for household income$_{2001}$ (gross earnings$_{2001}$) [hourly wage$_{2001}$] reverses from 11.0 (18.7) [16.1] to -18.1 (-11.2) [-10.9]. As such, conditioning on pre-exposure *trends* in our outcome measures might bias the results for the high versus medium comparison of the men strata.

Since we observe some regional clustering which might be related to regional authorities ('*Länder*') rather than the geographic regions that we control for in our main specification, we re-estimate all models using 15 '*Länder*' dummies instead the four regional dummies (Test F). Effect sizes remain similar for men. For women, they reduce (for several outcome measures) for the high versus low comparison while they increase for the high versus medium comparison. Still, however, earning gains for men remain well above the corresponding effects for women.

Next to these specification tests, we compare our findings with models estimated by nearest neighbor matching. In contrast to radius matching, nearest neighbor matching is less stable and the iterative approach as described in Section 4.3 does not converge. For reasons of completeness, we present the findings for nearest neighbor matching with some

---

[34] In order to test, how strongly an unmeasured variable must influence the selection process to undermine the implications of our study, we also considered the bounding approach originally proposed by Rosenbaum (2002). In general, the Rosenbaum-test does not appear to be very attractive, since it is a parametric test. Moreover, the implemented command in STATA (mhbounds, see Becker and Caliendo, 2007) was developed for binary outcomes and refers to the average treatment effects on the treated (i.e. ATETs) rather than ATEs which are in the focus of our study. Nevertheless, we experimented with the Rosenbaum-test for the men-strata where we find most of the significant effects in our study. For doing so, we recoded the continuous measures into binary outcomes by median-splitting the estimation samples in a first step. Of interest are the estimated Γ-values indicating for which size of a hidden bias the ATETs become insignificant with p > 0.1. In our study, Γ-values averaged over years for all significant ATEs (as reported in Figs. 4 and 5) measure 1.2 / 1.5 / 1.4 for net household income / gross earnings / net earnings when moving from medium to high SPE levels. The corresponding Γ-values when moving from low to high SPE levels are 1.8 / 1.9 for gross earnings / net earnings (since we did not find any significant ATE for net household income in the post exposure period, we cannot report any Γ-value for this comparison). The interpretation is as follows: our estimated effects on household income and earnings for men are insensitive to a hidden bias that would increase the odds by 20-90% of belonging to the treatment group, i.e. the group with comparably higher levels of SPE. While these Γ-values do not confirm that the CIA holds in our setting, they suggest that our main implications are not very sensitive to unobserved heterogeneity (detailed results are available upon request).

[35] To keep as many observations as possible, we only excluded individuals who were still in the education system in 2007 or later in the main specification.

[36] Moreover, this is confirmed by the common support statistics indicating comparably more observations off support (see Table B.4 in Appendix B) as well as the balancing statistics indicating comparably higher mean and medium biases for most covariates, in particular for the high versus low models (see Table B.4 in Appendix B).

[37] Since a considerably large "refreshment sample" of panelists was added in 2000 to the existing SOEP sample, we were not able to go further backwards. Otherwise, we would have lost nearly half of our estimation sample.

(Test G)[38] and without any iteration (Test H). However, since we are unable to restrict our analysis to *all* observations *on support* for a certain group, the sample reduction results of Lechner (2001) do not apply and we have no confidence in either of these results.

Finally, we compare our findings using the original 'pure' SPE measure instead of the radius measure as defined in Section 3 (Test I). Most ATEs remain positive. For instance, looking at the (H–M) models, the effects on gross earnings increase while they remain similar for net earnings. Only for income, they slightly reduce. Importantly, however, the standard errors of most estimates increase, partly substantially. We take this observation as suggestive evidence for the relevance of our radius measure in order to capture the individuals' willingness to travel for practicing sports. In other words: the results of the 'pure' measure are not precise and too noisy indicating that the treatment under this definition is not appropriate.

Summing up, while the estimated effects in Tests A to I differ in size as well as precision, our main finding remains, i.e. we observe a positive effect on net household income when moving from medium (partly also from low) to high SPE levels for both men and women. Moreover, the findings of our different robustness checks reinforce our interpretation that these effects are captured by earning gains for men rather than women living in the household. Admittedly, however, most estimates for outcomes other than household income and earnings are comparably noisy and volatile (which is also in line with our main findings discussed in Section 5.2).[39]

5.4. Effect heterogeneity

As mentioned previously, significant differences occur in terms of gender. While men and women exposed to comparably higher levels of SPE benefit from an increase in household income, additional analysis reveals that these gains in income are captured by earning gains for men rather than women living in the household. To determine whether this result is plausible, we need to more closely investigate men and women's differing sports participation habits.

To do so, we make use of the SOEP-Innovatiossample (SOEP-IS) (2019), which covers more than 3,000 households. It was established in 2012 (Richter and Schupp, 2015) and included a module on physical activity in 2013, 2015 and 2017.[40] According to these data, about 38 percent of women (35% of men) in the age cohort of interest (20–52 years) did *not* engage in any sports. Within our observation window (2013–2017), the share of inactive men remained rather stable, while the corresponding share of inactive women increased. As a consequence, the gender gap in terms of inactivity increased by about 5 percentage

points within four years. We also observe significant gender differences in the settings in which sport is practiced. In this regard, 16 (12) percent of women (men) engage in their first sport with a commercial sports provider. This is in line with the finding that 56 percent of members in fitness centers in Germany are female (Deloitte, 2018).

Overall, these figures suggest that women are generally less likely to engage in sports than men. Moreover, active women have a higher probability of engaging in sports outside any of the publicly funded sports facilities. Therefore, we confirm the plausibility of assuming that women have a lower probability of benefiting from higher levels of SPE than men.[41]

5.5. Exploring mechanisms

The underlying assumption of our analysis is that differences in per capita SPEs among muni-cipalities are associated with differences in the quantity and/or quality of sports facilities which translates into differences in LTPA and sports participation behavior of individuals.

Although we are unable to test the first part of this assumption at the local level given a lack of data, we find some supportive evidence at the regional level by correlating the average level of SPE with different indicators for sports facility provision in the 16 states (*Länder*).[42] Likewise, there is some correlational (and partly causal) evidence on the positive association between public expenditures and/or sports facility provision and LTPA (see Section 2). Whether SPE induced differences in the quantity and/or quality of sports facilities indeed influence sports participation behavior *in our setting*, is explored by Steckenleiter et al. (2019). They use the continuous nature of our SPE measure as well as the available sports participation measure in the SOEP for estimating dose-response functions. Result suggest, that moving from the low/medium to the medium/high SPE threshold, i.e. from around €20 to around €30, would merely translate to a change in at least monthly sports participation of 0.6 percentage points. As such, they conclude that higher levels of SPE do not increase the *probability to practice sports*.

It is important to note, however, that the available measure in the SOEP only considers the frequency of sports participation and does not measure any differences in intensity and/or duration of already active sports participants. A certain composition of all three characteristics describing sports participation is, according to the World Health Organization (2010), however, required for achieving positive health effects.[43] Therefore, we performed another more indirect test for differences in sports participation behavior by testing for differences in well-being and health. More precisely, we estimated and compared the effects for men between our main specification and the main specification including satisfaction with health, satisfaction with life and number of doctor visits per year, as well as the physical component (PCS) and the mental component summary scales (MCS) *from the corresponding year* as additional co-

---

[38] We stopped the process before any iteration for which the number of observations off support increased again.

[39] We also test whether panel attrition may be an issue by checking whether being in either of the SPE groups leads to selective attrition, which could potentially invalidate our estimates. To do so, we make use of the full data set after deleting all observations with missing values for our confounding variables and estimate the effects of SPE level on a binary outcome variable, which equals 1 if there is missing information in any outcome variable of interest of the respective year for a particular individual. Note that such a test goes beyond a pure panel attrition test and is somewhat stricter in that individuals also drop out of our analysis in years for which we do not observe any outcome (though they may still remain in the panel). Although only 1 of 10 estimates is significant at the 10 percent level for women when moving from medium to high SPE levels, 4 of 10 estimates for men and women are significant when moving from low to high SPE levels. Therefore, while we cannot fully rule out panel attrition bias in general, our test results indicate that the main findings for the medium versus high comparison are hardly affected (results are available upon request).

[40] The module, developed by Lechner and Pawlowski (2013), includes various questions about participation patterns in the first and second most frequently practiced sport. Unfortunately, we are not able to use these data in our main analysis because essential information (e.g., respondents' residence before 2013) is not available.

[41] We also tested whether any of the observed effects depend on age by further splitting the gender samples into two groups, i.e. men/women aged 20–36 years and men/women aged 37–52 years in 2001. We find mild evidence that the effects we found unfold for older rather than younger people. Likewise, we tested whether any labor market effects depend on whether children are living in the household by re-estimating our models for men/women living with and without children in the household separately. We find mild evidence that the effects we found unfold for people without rather than with children living in the household. Since, however, all sub sample estimations suffer from a considerable efficiency loss due to smaller sample sizes, we refrain from presenting and further discussing these results here (they are available upon request).

[42] We find some positive correlations between the average level of SPE and the total number of sports facilities as well as the area (in square meters per 1,000 inhabitants) covered by outdoor and indoor pools; the later correlation is significant at the 10 percent level (see Table D.1 in Appendix D).

[43] Supported by various studies offering empirical evidence for this, the World Health Organization's guidelines suggest that 150 min (75 min) per week of moderate (vigorous) regular physical activity is required to achieve positive health effects.

variates in the selection process. The idea is that any labor market effects caused by an increase in health-enhancing sports participation should disappear when controlling for differences in health status.[44] Figure D.1 in Appendix D displays the effects of different levels of SPE on household income, earnings, working time, wages and full-time employment for *men* during the post-exposure period excluding and including our health measures from the corresponding year as additional covariates in the selection process. Visual inspection of these graphs suggests that some effects indeed decreased when controlling for the various health measures. This reduction is more striking for even years (2008, 2010 and partly 2012), for which we could control for the most sophisticated health measures available, the PCS and MCS.

Table D.2 in Appendix D quantifies the overall effect of controlling for the health measures. The table displays the ATEs for both specifications (ATE and ATE$_{Health}$) and the average change (in percentage) between ATE and ATE$_{Health}$ calculated for all years in the period 2007–2012 (Change^1) as well as only for years with significant estimates (Change^2).[45] After controlling for the different health measures, we found a reduced effect for net household income and earnings of approximately 15 percent as well as a reduced effect for wage (being employed full time) of approximately 8–10 (6) percent.[46] Importantly, when controlling for PCS and MCS is possible, these values increase to approximately 40 percent for income and earnings and approximately 20 percent for wage and being employed full time.

Overall, these findings suggest that the estimated positive effects on labor market outcomes are (at least) partly driven by health-enhancing sports participation. This interpretation is reinforced by additional results from the *mean* model estimates. For instance, while the *average* gain in net household income when moving from medium to high SPE levels is reduced for men by approximately 12 percent, it hardly changes for women. Moreover, we do not observe any other significant ATE anymore. As such, we conclude, that while Steckenleiter et al. (2019) do not confirm that higher SPEs *make more people active*, our indirect test suggests, that higher SPEs may *increase activity levels of already active individuals* in order to meet (or at least come closer to) the threshold as defined by the World Health Organization and thus benefit from health gains which are rewarded in the labor market. This is in line with Lechner and Sari (2015) who find a positive income effect for an activity increase of already active individuals while the change from inactivity to only a moderate level of sports and exercise was found to be too small to generate such effects.

At the same time, however, these findings also suggest that other channels beyond health remain relevant. In this regard, Lechner (2009) notes that such channels might relate to, for example, social networking. Indeed, using SOEP data, Schüttoff et al. (2018) show that regular sports participation positively affects adolescents' social capital through volunteering, helping friends and civic involvement. Since information on these measures is not available for 2008, 2010 and 2012, any joint effect analysis together with PCS and MCS is impossible. However, results from the *mean* models which condition on the three social capital indicators in the post exposure period suggest that social capital might be a relevant channel. While most effects are considerably smaller, the gain in household income for women when moving from medium to high SPE levels is the only significant ATE that remains.

Overall, while the estimated household income effect when moving from medium to high SPE levels substantially reduces for men (i.e. from €153 to €136 / €71) when additionally controlling for health /

social capital indicators in the post exposure period, the corresponding effect for women remains remarkably stable (i.e. €157 / €160 / €159). As such, these findings reinforce our interpretation about the household income effect being captured by earning gains for (active) men rather than women living in the household.

Finally, since the (testable) individual gains in social capital and health do not fully explain the effects we find, some other (institutional) channels might also be relevant. However, data to test, for instance, whether well-equipped sports infrastructure serves as a soft factor in attracting firms or whether SPE-induced earning gains exist for individuals working in local firms contracted for sports facility-related renovation, operation or maintenance work, are not available.[47]

6. Conclusion

By merging administrative data of all municipalities in Germany with individual data from the SOEP, we explore whether local public expenditures on sports facilities influence individual labor market outcomes. In order to approach the different endogeneity issues in this setting, we exploit the panel structure of our data and control for confounding influences at the *aggregate* (municipality) and *disaggregate* (individual) levels. Our analysis reveals sizable effects of public expenditures on individual labor market outcomes. For both men and women, we observe a significant (average) increase in household monthly net income of approximately €150 (approximately 5.8% of average household income[48]) when moving from medium (€20–€31) to high (€31–€85) levels of SPE. For a medium-sized city with 75,000 inhabitants, this translates into an average increase in annual spending of approximately €2.5 million. Assuming a 6 percent discount rate like Jackson et al. (2015), this corresponds to total investment with a present value of around €11 million over five years. Interestingly, the effects found in this study are captured by earning gains for men rather than women living in the household. Additional analysis confirms and supports these apparent gender differences. Moreover, we find some suggestive evidence that improved well-being and health are possible mechanisms.

With this study, we contribute to the rich literature analyzing the effects associated with public expenditures, which to date has largely neglected investigating types of public spending that might *indirectly* improve health, educational or labor market outcomes. However, while the informative micro-level data available enables us to provide the first study exploring the *long-term* effects of *local* public sports facility expenditures on *individual* labor market outcomes, some limitations must be noted directing to some relevant avenues for future research.

First, even though the data we used is rich, sample size issues prevent us from controlling for differences in medium- and long-term pre-exposure trends in labor market outcomes between low, medium, and high spending municipalities. Likewise, we are unable to comprehensively account for spatial heterogeneity in our models. As such, even though the tests we implemented in this regard (i.e. Test E and Test F) do not cast any doubts on our main findings, concerns about the validity of these (untestable) identifying assumptions remain. *Second*, and related to the aforementioned point, we are unable to explore any spa-

---

[44] Note that data for constructing the PCS and MCS measures were only gathered in even years since 2002 and therefore are not available for 2007, 2009 and 2011.

[45] Both change indices were only calculated if either ATE or ATE$_{Health}$ is significant for at least two years.

[46] We observe too few significant effects for calculating the corresponding changes for working time.

[47] According to a report by the Federal Ministry of Economic Affairs and Energy (2012), the biggest share of money for value-increasing measures and the construction of sport facilities is spent for contracts with specialized firms in structural and civil engineering. It seems plausible to assume that only few firms in structural and civil engineering exist that are specialized on sports facilities. Therefore, it is unlikely that contracting such firms directly unfolds positive effects on the local labor market. At the same time, however, considerable shares of operation and investment costs are also caused by contracting smaller (and eventually local) firms for preparatory construction work, installations (e.g. by plumbers), or others. As such, there might be indeed some effects on individual labor market outcomes by stimulating economic activities of such local firms.

[48] Average household net income is calculated as frequency weighted average of the six income figures for 2007, as displayed in Table 2.

tial dependency, i.e. the degree to which the decisions to invest in sports facilities for a municipality depend on the decisions of the neighboring municipalities. *Third*, further data restrictions and availabilities prevent us from testing all plausible mechanisms in a robust way. In this regard, testing whether well-equipped sports infrastructure might indeed serve as a soft factor in attracting firms or whether and to what extent multiplier effects occur would be interesting per se. Moreover, any such evidence could add further credibility to the comparably large effect sizes we find in our setting. *Fourth*, in contrast with Jackson et al. (2016) and others who examine the effects of local school spending, there is no clear guidance on how to specify the exposure period in our setting. Although our findings suggest that controlling for exposure to different levels of SPE over five years is sufficient to reveal some positive effects, specifying more precisely the conditions under which these effects unfold requires more attention in future studies. *Fifth*, even the more established literature from settings such as health, education and labor markets remains rather silent about opportunity cost arguments to date. Since the budget for sports is frequently in competition for scarce public money available for other 'cultural' areas such as museums or theaters, it seems relevant to explore whether and to what extent public spending on other areas might be associated with similar effects as found for sports in this study. However, the identifying assumptions are expected to differ considerably compared to the setting analyzed in this study (see Fig. 3) since confounding issues and mechanisms are different. Moreover, whether this will be possible in the future depends not only on data quality but also on data protection rules and regulations, which have also considerably limited the scope for implementing more comprehensive econometric techniques in our study.

Supplementary materials

Supplementary material associated with this article can be found, in the online version, at doi:10.1016/j.labeco.2021.101996.

References

Ahlfeldt, G.M., Kavetsos, G., 2014. Form or function? The effect of new sports stadia on property prices in London. J. R. Statistical Society 177 (1), 169–190.

Alba, R.D., Logan, J.R., 1991. Variations on two themes: racial and ethnic patterns in the attainment of suburban residence. Demography 28 (3), 431–453.

Andersen, H.A., Mühlbacher, A., Nübling, M., Schupp, J., Wagner, G.G., 2007. Computation of standard values for physical and mental health scales scores using the SOEP version of SF-12v2. Schmollers Jahrb. 127 (1), 171–182.

BIK – Aschpurwis & Behrens GmbH (2001). BIK-Regionen. Ballungsräume, Stadtzentren, Mittel-/Unterzentrengebiete. Methodenbeschreibung zur Aktualisierung [BIK regions. Agglomeration, city centers, medium and small regional centers]. Retrieved October 2, 2018 http://www.bik-gmbh.de/texte/BIK-Regionen2000.pdf

Becker, G.S., 1965. A theory of the allocation of time. Econ. J. 75 (299), 493–517.

Becker, S.O., Caliendo, M., 2007. Sensitiv analysis for average treatment effects. Stata J. 7 (1), 71–83.

Black, N., Johnston, D.W., Propper, C., Shields, M.A., 2019. The effects of school sports facilities on physical activity, health and socioeconomic status during adulthood. Soc. Sci. Med. 220, 120–128.

Bodory, H., Camponovo, L., Huber, M., Lechner, M., 2018. The finite sample performance of inference methods for propensity score matching and weighting estimators. J. Bus. Econ. Stat. doi:10.1080/07350015.2018.1476247.

Brink, L.A., Nigg, C.R., Lampe, S.M.R., Kingston, B.A., Mootz, A.L., van Vliet, W, 2010. Influence of schoolyard renovations on children's physical activity: the learning landscapes program. Am. J. Public Health 100 (9), 1672–1678.

Busemeyer, M., 2007. Determinants of public education spending in 21 OECD democracies, 1980–2001. J. Eur. Public Policy 14 (4), 582–610.

Cabane, C., Lechner, M., 2015. Physical activity of adults: A survey of correlates, determinants, and effects. J. Econ. Stat. (Jahrb. Nationalökonomie Stat.) 235 (4-5), 367–402.

Charness, G., Gneezy, U., 2009. Incentives to exercise. Econometrica 77 (3), 909–931.

Cornelissen, T., Pfeifer, C., 2008. Sport und Arbeitseinkommen – Individuelle Ertragsraten von Sportaktiven in Deutschland [Sport and earnings in Germany]. Rev. Econ. 59 (3), 244–255.

Cremieux, P., Ouellette, P., Pilon, C., 1999. Health care spending as determinants of health outcomes. Health Econ. 8 (7), 627–639.

Dallmeyer, S., Wicker, P., Breuer, C., 2017. Public expenditure and sport participation: an examination of direct, spillover, and substitution effects. Int. J. Sport Finance 12 (3), 244–264.

Dell, M., Jones, B.F., Olken, B.A., 2014. What do we learn from the weather? The new climate–economy literature. J. Econ. Lit. 52 (3), 740–798.

Deloitte (2018). Der deutsche Fitnessmarkt [The German fitness market]. Benchmarking report.

Downward, P., Dawson, A., Dejonghe, T., 2009. Sports Economics. Theory, Evidence and Policy. Elsevier, Amsterdam et al..

Downward, P, Frick, B., Humphreys, B.R., Pawlowski, T., Ruseski, J., Soebbing, B.P., 2019. The SAGE Handbook of Sports Economics. SAGE Publications Ltd, Thousand Oaks (California).

DSSV, 2007. Eckdaten 2007 der deutschen Fitnesswirtschaft [Core figures about the German fitness market]. Deutscher Sportstudio Verband e.V, Hamburg.

Eccles, J.S., Harold, R.D., 1991. Gender differences in sport involvement: applying the Eccles' expectancy-value model. J. Appl. Sport Psychol. 3 (1), 7–35.

Federal Ministry of Economic Affairs and Energy (2012). Die wirtschaftliche Bedeutung des Sportstättenbaus und ihr Anteil an einem zukünftigen Sportsatellitenkonto [The economic significance of sports facilities and their share in a sports satellite account] Retrieved on February 17, 2020 https://www.bmwi.de/Redaktion/DE/Publikationen/Studien/abschlussbericht -sportstaettenbau.pdf?_blob=publicationFile&v=7

Federal Ministry of Economic Affairs and Energy (2019). Sport inner- oder außerhalb des Sportvereins: Sportaktivität und Sportkonsum nach Organisationsform [Sports in sport clubs and beyond: sports participation and consumption patterns by organizational format] Retrieved on January 20, 2021 https://sportsatellitenkonto.de/wp-content/uploads/2019/12/16_BISp-Sport-inner-oder-au%C3%9Ferhalb-des-Sportvereins-2019.pdf

Federal Ministry of Interior (2010). Zwölfter Sportbericht der Bundesregierung [12th Sports report for the Federal Government]. Berlin.

Federal Statistical Office (2016). Rechnungsergebnisse der öffentlichen Haushalte für soziale Sicherung und für Gesundheit, Sport, Erholung –2010 [Annual financial statement of the public authorities for social security, health, sport, recreation]. Wiesbaden: Statistisches Bundesamt.

Federal Statistical Office (2017). Verbraucherpreisindex (inkl. Veränderungsraten) in Deutschland, Code: 61111-0001, Stand: 2. November 2017 [Consumer price index (including change rates) in Germany, Code: 61111-0001, as of November 2, 2017]. Retrieved November 2, 2017 https://www.genesis.destatis.de/genesis/online?sequenz=statistikTabellen&selectionname=61111

Federal Statistical Office, 2019. Volkswirtschaftliche Gesamtrechnungen der Länder VGRdL – Einwohner 1991 bis 2018 [Regional accounts VGRdL – Inhabitants 1991 to 2018]. Statistisches Bundesamt, Wiesbaden.

Fitzenberger, B., Schnabel, R., Wunderlich, G., 2004. The gender gap in labour market participation and employment: a cohort analysis for West Germany. J. Popul. Econ. 17 (1), 83–116.

Foremny, D., Riedel, N., 2014. Business taxes and electoral cycle. J. Public Econ. 115, 48–61.

Fuest, C., Peichl, A., Siegloch, S., 2018. Do higher corporate taxes reduce wages? Micro evidence from Germany. Am. Econ. Rev. 108 (2), 393–418.

Gauss, P., Riedel, N. & Simmler, M. (2019). The incidence of local public goods on wages: evidence from Germany. Working Paper.

German Olympic Sports Confederation (2019). Integration through sport. Retrieved April 29, 2016 https://cdn.dosb.de/user_upload/www.integration-durch-sport.de/Service/Info-Material/Flyer_Programm_ENGLISCH.pdf.

German Statutory Pension Insurance Scheme (Deutscher Rentenversicherung Bund) (2017). Rentenversicherung in Zeitreihen [Time series of the German pension insurance scheme]. Berlin.

Goebel, J., Grabka, M.M., Liebig, S., Kroh, M., Richter, D., Schröder, C., Schupp, J., 2018. The German Socio-Economic Panel (SOEP). Jahrb. Nationalökonomie Stat./J. Econ. Stat. doi:10.1515/jbnst-2018-0022, (online first).

Ham, J., Swenson, C., İmrohoroğlu, A., Song, H., 2011. Government programs can improve local labor markets: evidence from state enterprise zones, federal empowerment zones and federal enterprise community. J. Public Econ. 95 (7-8), 779–797.

Harris, A., Evans, W., Schwab, R., 2001. Education spending in an aging America. J. Public Econ. 81 (3), 449–472.

Hoehner, C.M., Ramirez, L.K.B., Elliott, M.B., Handy, S.L., Brownson, R.C, 2005. Perceived and objective environmental measures and physical activity among urban adults. Am. J. Prev. Med. 28 (2), 105–116.

Huang, H., Humphreys, B.R., 2012. Sports participation and happiness: evidence from US microdata. J. Econ. Psychol. 33 (4), 776–793.

Huber, M., Lechner, M., Steinmayr, A., 2015. Radius matching on the propensity score with bias adjustment: tuning parameters and finite sample behaviour. Empir. Econ. 49 (1), 1–31.

Huber, M., Lechner, M., Wunsch, C., 2013. The performance of estimators based on the propensity score. J. Econom. 175 (1), 1–21.

Humphreys, B.R., Ruseski, J.E., 2007. Participation in physical activity and government spending on parks and recreation. Contemp. Econ. Policy 25 (4), 538–552.

Imbens, G.W., 2004. Nonparametric estimation of average treatment effects under exogeneity: a review. Rev. Econ. Stat. 86, 4–29.

Jackson, C.K., Johnson, R., Persico, C., 2016. The effects of school spending on educational and economic outcomes: evidence from school finance reforms. Q. J. Econ. 131 (1), 157–218.

Kokolakakis, T., Lera-López, F., Castellanos, P., 2014. Regional differences in sports participation: the case of local authorities in England. Int. J. Sport Finance 9 (2), 149–171.

Lechner, M., 2001. Identification and estimation of causal effects of multiple treatments under the conditional independence assumption. In: Lechner, M., Pfeiffer, F. (Eds.), Econometric Evaluation of Labour Market Policies. Physica, Heidelberg, pp. 43–58.

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

Lechner, M., 2015. Sports, exercise, and labor market outcomes. increasing participation in sports and exercise can boost productivity and earnings. In: IZA World Labor, 126, pp. 1–10.

Lechner, M., Downward, P., 2017. Heterogeneous sports participation and labour market outcomes in England. Appl. Econ. 49 (4), 335–348.

Lechner, M., Miquel, R., Wunsch, C., 2011. Long-run effects of public sector sponsored training in West Germany. J. Eur. Econ. Assoc. 9 (4), 742–784.

Lechner, M. & Pawlowski, T. (2013). On the socio-economic effects of physical activity. Unpublished SOEP-IS-2013 proposal (accepted).

Lechner, M., Sari, N., 2015. Labor market effects of sports and exercise: evidence from Canadian panel data. Labour Econ. 35, 1–15.

Lichter, D.T., LeClere, F.B., McLaughlin, D.K., 1991. Local marriage markets and the marital behavior of black and white women. Am. J. Sociol. 96 (4), 843–867.

Pawlowski, T., Breuer, C., 2012. Die Finanzpolitische Bedeutung des Sports in Deutschland (Fiscal Political Importance of Sports in Germany). Springer-Gabler Research, Wiesbaden.

Porter, M.E., 1998. Clusters and the new economics of competition. Harv. Bus. Rev. 76, 77–90.

Richter, D., Schupp, J., 2015. The SOEP innovation sample (SOEP IS). Schmollers Jahrb. 135 (3), 389–399.

RDC of the Federal Statistical Office and Statistical Offices of the Länder (2018a). Public finance statistics [Jahresrechnungsstatistik der öffentlichen Haushalte], survey years 2001–2006, own calculations.

RDC of the Federal Statistical Office and Statistical Offices of the Länder (2018b). Annual balance sheets of public funds, institutions and enterprises [Jahresabschlüsse der öffentlichen Fonds, Einrichtungen und Unternehmen], survey years 2001–2006, own calculations.

Rosenbaum, P.R., 2002. Observational studies (2nd edition). Springer, New York.

Schüttoff, U., Pawlowski, T., Downward, P., Lechner, M., 2018. Sports participation and social capital formation during adolescence. Soc. Sci. Q. doi:10.1111/ssqu.12453.

Socio-Economic Panel (SOEP), 2014. Data for years 1984-2012. version 29. SOEP doi:10.5684/soep.v29.

SOEP-Innovationssample (SOEP-IS), 2019. Data for years 1998-2017. SOEP doi:10.5684/soep.is.2017.

Sports Minister Conference (2003). Sportstättenstatistik der Länder [Statistics on sports facilities of the states]. Berlin.

Steckenleiter, C., Lechner, M., Pawlowski, T., Schüttoff, U., 2019. Do local public expenditures on sports facilities affect sports participation in Germany? Economics Working Paper Series 1905. University of St. Gallen, School of Economics and Political Science.

Steinmayr, A., Felfe, C., Lechner, M., 2011. The closer the sportier? Children's sports activity and their distance to sports facilities. Eur. Rev. Aging Phys. Act. 8, 67–82.

Stevenson, B., 2010. Beyond the classroom: using title IX to measure the return to high school sport. Rev. Econ. Stat. 92 (2), 284–301.

Tiebout, C., 1956. A pure theory of local expenditures. J. Polit. Econ. 64 (5), 416–424.

Warburton, D., Nicol, C.W., Bredin, S.S.D, 2006. Health benefits of physical activity: the evidence. Can. Med. Assoc. J. 174 (6), 801–809.

Wicker, P., Breuer, C., Pawlowski, T., 2009. Promoting sport for all to age-specific target groups: the impact of sport infrastructure. Eur. Sport Manage. Q. 9 (2), 103–118.

Wicker, P., Hallmann, K., Breuer, C., 2013. Analyzing the impact of sport infrastructure on sport participation using geo-coded data: evidence from multi-level models. Sport Manage. Rev. 16 (1), 54–67.

World Health Organization (2009). Promoting sport and enhancing health in European Union countries: a policy content analysis to support action. Retrieved January 19, 2021 https://www.euro.who.int/_data/assets/pdf_file/0006/147237/e95168.pdf

World Health Organization (2010). Global recommendations on physical activity for health. Retrieved July 14, 2016 http://www.who.int/dietphysicalactivity/publications/9789241599979/en/index.html.

World Health Organization (2013). Global Action Plan for the Prevention and Control of Noncommunicable Diseases 2013-2020. Retrieved September 12, 2018 https://www.who.int/nmh/events/ncd_action_plan/en/.