Universität St.Gallen

Do local public expenditures on sports facilities affect sports participation in Germany?

Carina Steckenleiter, Michael Lechner, Tim Pawlowski, Ute Schüttoff

April 2019 Discussion Paper no. 2019-05

School of Economics and Political Science,
Department of Economics

University of St.Gallen

Editor: Vanessa Pischulti
University of St.Gallen
School of Economics and Political Science
Department of Economics
Müller-Friedberg-Strasse 6/8
CH-9000 St.Gallen
Phone +41 71 224 23 07
Email seps@unisg.ch

Publisher: School of Economics and Political Science
Department of Economics
University of St.Gallen
Müller-Friedberg-Strasse 6/8
CH-9000 St.Gallen
Phone +41 71 224 23 07

Electronic Publication: http://www.seps.unisg.ch

Do local public expenditures on sports facilities affect sports participation in Germany? [1]

Carina Steckenleiter, Michael Lechner[2], Tim Pawlowski[3], Ute Schüttoff

Author's address: Carina Steckenleiter
Swiss Institute for Empirical Economic Research (SEW)
University of St.Gallen
Varnbüelstrasse 14
CH-9000 St.Gallen
Email carina.steckenleiter@unisg.ch

Author's address: Michael Lechner
Swiss Institute for Empirical Economic Research (SEW)
University of St.Gallen
Varnbüelstrasse 14
CH-9000 St.Gallen
Email michael.lechner@unisg.ch
Website michael-lechner.eu

Author's address: Tim Pawlowski
University of Tübingen
Wilhelmstrasse 124
DE-72074 Tübingen
Email tim.pawlowski@uni-tuebingen.de

Author's address: Ute Schüttoff
University of Tübingen
Wilhelmstrasse 124
DE-72074 Tübingen
Email ute.schuettoff@uni-tuebingen.de

---

[1] Support from the German Research Foundation (PA 2500/1-1) and the Swiss National Science Foundation (100018E-161598) is gratefully acknowledged. We thank SOEP for providing data access at the SOEP Research Data Center at DIW Berlin and Jan Goebel for providing expertise on SOEP regional data. Moreover, we thank Steffi Dierks, Matthias Klumpe and Steffen Lauf (Statistical Office for Berlin-Brandenburg, Department: Microdata, Analysis and Research Data Center) as well as Kristin Nowak (Research Data Center of the Statistical Offices of the Länder, Office Stuttgart) for providing data access and comprehensive support with regard to the 'Public Finance Statistics' as well as the 'Annual Balance Sheets of Public Funds, Institutions and Enterprises'. We are indebted to Tim Wallrafen for his help in collecting, cleaning and transforming the regional data. The usual disclaimer applies.

[2] Michael Lechner is also affiliated with CEPR, London, CESIfo, Munich, IAB, Nuremberg, IZA, Bonn, and RWI, Essen.

[3] Tim Pawlowski is also affiliated with the LEAD Graduate School and Research Network.

Abstract

This paper analyzes the effect of local public expenditures on sports facilities on sports participation in Germany. To this end, we construct a new database containing public expenditures at the municipality level and link this information with individual level data. We form locally weighted averages of expenditures based on geographic distances since people also benefit from expenditures of neighboring municipalities. We analyze how effects of sports facility expenditures change with different expenditures levels ("dose-response relationship") and find no effect of local public expenditures on sports facilities on the probability to practice sports. These findings are robust across different age groups and municipality sizes.

Keywords

Sports, local public sports expenditure, continuous treatment, treatment effects dose-response surface, semiparametric estimation

JEL Classification

Z20, H72, H75, C21

1 Introduction

Supporting individual physical activity is an important goal for many governments. In that regard, countries around the world spent substantial amounts of money to foster sports participation. For example, Pawlowski and Breuer (2012) estimate the total amount of annual sports related public expenditures in Germany at around 10 billion EUR. Similar to other public expenses, public expenditures in the area of sports are reasoned with the existence of market failures as well as public and merit good characteristics (Langer, 2006). One line of argument is that sports participation is associated with a reduction of health care costs through improved mental and physical well-being as well as the accumulation of social and human capital (Federal Ministry of Interior, 2010).

Despite this positive association with physical activity, recent research suggests that insufficient physical activity has become more prevalent in high-income countries over the last 15 years (36.8% in 2016) and was more than twice as high as in low income countries in 2016 (Guthold, Stevens, Riley, & Bull, 2018).[1] Since physical inactivity has been shown to increase the risk of suffering from Noncommunicable Diseases (NCD)[2] and NCDs are the leading cause of death across the globe (World Health Organization, 2013), increasing physical activity levels is defined as one key component in the WHO's *"Global Action Plan for the Prevention and Control of Noncommunicable Diseases (NCD) 2013-2020"*. Construction and maintenance of sports infrastructure is one item on the list of policy options formulated by the WHO to push the reduction of physical inactivity forward.[3]

[1] Guthold et al. (2018) define insufficient physical activity as *"adults not meeting the WHO recommendations' on physical activity for health - ie, at least 150 min of moderate-intensity, or 75 min of vigorous-intensity physical activity per week, or any equivalent combination of the two."*(Guthold et al., 2018, p. e1078).

[2] NCDs encompass cardiovascular diseases, diabetes, cancers as well as respiratory conditions (World Health Organization, 2018).

[3] The precise text is as follows: *"Creation and preservation of built and natural environments which support*

Despite the large amount of money spent to supply sports facilities (e.g., German municipalities spent around 4.3 Billion EUR in 2006)[4], little is known whether these public expenditures are effective in fostering sports participation. Only few studies have investigated the link between public sports expenditures and physical activity *in general*. Even less is known about the link of facility expenditures and sports participation *in particular*.[5] Kokolakakis, Lera-López, and Castellanos (2014) analyse physical activity at the level of local authorities in England and find no correlation of sports expenditures with sports participation. Downward and Rasciute (2015) find that the intensity of female sports participation is positively associated with satisfaction of the provided facilities in a study using English data. Humphreys and Ruseski (2007) investigate the effect of public funding on leisure sports in the US. They find that higher public expenditures on parks go along with higher outdoor activities. Similarly, Dallmeyer, Wicker, and Breuer (2017) look at the relationship of different types of sports expenditures and sports participation with the SOEP. They report positive associations between sports facilities expenses at the regional (state) level and sports participation for Germany.[6] This brief overview suggests that evidence regarding the effect of local sports facility expenditures on sports participation is generally scarce. Moreover, the

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." (World Health Organization, 2013, p. 34).

[4]RDC of the Federal Statistical Office and Statistical Offices of the Länder (2018a, 2018b), own calculations.

[5]In a related study, Pawlowski, Steckenleiter, Wallrafen, and Lechner (2019) analyse the link of sports facility expenditures and individual labor market outcomes using the same data set as this study. Their results suggest that high levels of sports expenditures impact earnings for men positively. They also find that women are more likely to be active in other settings than in sports facilities that are publicly funded.

[6]Related to our research question, a set of studies investigated the link of the availability of sports infrastructure and sports participation for Germany (e.g. Reimers et al., 2014; Steinmayr, Felfe, & Lechner, 2011; Wicker, Breuer, & Pawlowski, 2009). Wicker et al. (2009) emphasize that different types of sports infrastructure are relevant for different age groups. Steinmayr et al. (2011) and Reimers et al. (2014) investigate the impact of proximity of sports infrastructure on sports participation for children. They find that distance to sport facilities is of importance (for girls) in rural areas.

aforementioned studies exhibit shortcomings in some dimensions.

A general challenge for such analysis is that expenditures on sports facilities are not randomly assigned and people might choose a municipality based on the provision of certain infrastructure such as childcare, public transport or quality of sports facilities. This argument was already made by Tiebout (1956) who notes that a person chooses the municipality *"which best satisfies his preference pattern for public goods"* (Tiebout, 1956, p. 418). Second, there is no evidence how effects of sports expenditures change with different expenditure levels (the so called "dose-response relationship"). Gaining knowledge about the dose-response relationship is however important for policy makers to allocate tax-payer money cost-efficiently. Exemplarily, expenditures on sports facilities could potentially impact sports participation only once a certain level of expenditures has been reached. It may also be that no additional gains arise from increasing expenditures beyond a certain level. Third, there is considerable heterogeneity concerning the type of expenditures included in the sports expenditure measures used. Only few studies are explicit about this and focus on a clearly defined expenditure category. Finally, expenditure data used in earlier studies was only available on a very aggregated level. Since the provision of sports facilities (e.g. gyms, swimming-pools) is largely in the hand of local authorities in the case of Germany, it is important, however, to conduct the analysis on a fine regional level.[7]

We try to overcome the aforementioned shortcomings by constructing a new data base with public sports expenditures for Germany at the level of municipalities. Municipalities represent the lowest administrative level in Germany, we thus measure expenditures on the finest regional level possible. We also account for the fact that individuals also benefit from

[7]This is acknowledged, for instance, by Humphreys and Ruseski (2007) who note that deriving policy conclusions from their study is difficult since their expenditure measure is too aggregated.

expenditures of neighboring municipalities (although less due to higher travel costs). To do so, we form locally weighted averages (based on geographic distances) of public sports expenditures. Apart from expenditures on sports facilities, people potentially benefit from spending of neighboring municipalities in other areas of public expenditures as well. Hence, our way of forming distance weighted expenditure averages might also be useful for evaluations of public spending in other areas.

The expenditure data is supplemented with several other local characteristics and merged with individual level data from the German Socio-Economic Panel (SOEP), which is a widely used panel study for Germany (Wagner, Frick, & Schupp, 2007). Given the information available in this merged data base, we argue that a selection-on-observables assumption is a credible way to identify the effect of local expenditures on individual sports participation. Subsequently, we estimate so-called dose-response relationships. Compared to the binary treatment case, empirical studies which use propensity score methods for the continuous treatment case are relatively rare so far.[8] In this paper, we use the approach proposed by Flores et al. (2012), since it allows us to estimate the dose-response function semi-parametrically in a very flexible and computationally efficient way.[9] One practical issue which arises upon implementation is how balancing can be effectively assessed in a continuous treatment framework. The purpose of balancing tests is to assess whether one was successful in creating groups that are comparable with respect to their characteristics. We provide a possible alternative approach to assess balancing which might be useful for other

[8]Starting from Imbens (2000) and Hirano and Imbens (2004), Imai and Van Dyk (2004), Flores, Flores-Lagunes, Gonzalez, and Neumann (2012), Kluve, Schneider, Uhlendorff, and Zhao (2012), Moodie and Stephens (2012), Dammert and Galdo (2013), Bia, Flores, Flores-Lagunes, and Mattei (2014), Choe, Flores-Lagunes, and Lee (2015) and Wu, Mealli, Kioumourtzoglou, Dominici, and Braun (2018) have developed or applied such methods for continuous treatments.

[9]Computational efficiency is very important since for reasons of data protection all estimations had to be carried out at the SOEP Research Data Center in Berlin.

researchers who consider using these methodologies in applied work.

Overall, we find no effect of expenditures on sports facilities on the probability to practise sports. This result sheds considerable doubt on the main justification for such expenditures, namely that they increase public health by reducing the level of inactivity in society. Our results are robust across different age groups, measuring sports participation at different points in time, the inclusion of lagged activity and expenditure information as well as splitting the sample into smaller and larger municipalities.

The next section introduces the new regional-individual data base and presents relevant descriptive statistics. This is followed by introducing the econometric framework in Section 3. Results as well as robustness checks are presented in Section 4. Sections 5 concludes and identifies avenues for future research. Appendices A and B contain more information on the data and descriptive statistics. Appendix C presents further figures and tables with regard to GPS-estimation, common support and balancing tests, Appendix D contains robustness checks and Appendix E additional estimation results.

2 Data

2.1 Regional expenditure data

Our data entails administrative expenditure data on the municipality level from 2001 to 2006 (RDC of the Federal Statistical Office and Statistical Offices of the Länder, 2018b).[10] [11] Our measure of relevant local public expenditures is the sum of net expenditures on sports

[10] We cannot use more recent years since accounting standards across German states became too different from 2007 onwards.

[11] Pawlowski et al. (2019) use the same data analyzing the impact of expenditures on sports facilities on individual labor market outcomes. Therefore, the sections describing the data set and the sports expenditure

facilities (incl. swimming pools) on the municipality level.[12] Expenses in the area of sports facilities encompass construction and maintenance expenses for various types of sports facilities (Federal Statistical Office, 2016). Since many municipalities outsource sports facilities and swimming-pools into municipal companies (and these numbers are not contained in the accounting data of the municipalities), we also collected expenditure data for municipal companies and integrated them into our data base (RDC of the Federal Statistical Office and Statistical Offices of the Länder, 2018a).[13] The expenditure data is combined with regional information from the so-called INKAR database,[14] population and district data (Federal Statistical Office, 2011) as well as meteorological data over a longer period of time (DWD (Deutscher Wetterdienst), 2016). All data were merged over the municipality identifier on a year-by-year basis.[15]

Since individuals also benefit from expenditures of neighboring municipalities, we construct distance weighted expenditures measures where expenditures of neighboring municipalities are also accounted for. In this section, we explain how we define the relevant vicinity of communities and also how we assign weights. Pawlowski, Breuer, Wicker, and Poupaux (2009) find that the average maximum willingness to travel to a sports facility is about 26 minutes for the Stuttgart region. Since travel speed varies across urban and rural areas,

measure naturally have a high degree of overlap in both papers. However, this is indispensable in order to fully understand the empirical strategies employed in each paper.

[12]Due to municipality mergers in particular in the new federal states of the former East Germany (the number of municipalities decreased from 12,930 in 2002 to 12,291 in 2006) substantial manual preparation was necessary to harmonize the data across different years as well as with other data sources.

[13]These data had to be accessed on-site at one of the regional data centers of the Federal Statistical Office. A detailed description of all steps of data preparation can be found in Pawlowski et al. (2019).

[14]Indicators and maps on spatial and urban development in Germany (INKAR), 2013 edition, are made available by the BBSR (Bundesinstitut für Bau-, Stadt- und Raumforschung) (2013) of the Federal Office for Building and Regional Planning (BBR), http://www.inkar.de.

[15]Not all data exist on the municipal level. For meteorological data, we assign information from the closest weather station to each municipality. For a part of the municipalities, INKAR-data was only available on the level of municipal associations. In these cases, we assign municipalities the data from the municipal association it belongs to.

we assign municipalities to three different categories based on INKAR-information: *"core cities"* (22% of municipalities in 2006), *"densely populated suburban districts"* (45%) and *"rural areas"* (33%). Information from the German mobility report (infas, DLR, 2010) is used to calculate how far one is able to travel within 26 minutes.[16] Information from the report translates to travel distances of 11.3 km (core cities), 13.4 km (densely populated suburban districts) and 14.3 km (rural areas). Hence, depending on the category a municipality belongs to, municipalities in the corresponding distance are accounted for in the sports expenditure measure. This means, for instance, that for rural communities all municipalities within a distance of 14.3 km are taken into account.[17]

Per capita sports facility expenditures in a given year are defined as follows:

$$\text{sports-exp}_c = \frac{x}{r}\ \text{sports-exp-pc}_c + \left\{ \frac{(r-x)}{r} \left[\frac{\sum w_i'\ \text{sports-exp-neighbor}_i}{\sum w_i'\ \text{Population}_i} \right] \right\} \tag{1}$$

where x is the distance between the centroid of the closest neighboring municipality and the centroid of the municipality itself. r reflects the radius taken into account and is depending on the category of the municipality either 11.3 km, 13.4 km or 14.3 km. Hence, the larger the respective municipality itself, the less weight the neighboring municipalities get. For example, if the centroid of the closest neighboring municipality is 7 km and the municipality is a core city (hence r is 11.3 km), per-capita expenditures of the municipality itself enter with a weight of 7/11.3 km and the neighboring municipalities with 4.3/11.3

[16] Average travel speed in core cities is indicated with 26 km per hour. Correspondingly, it is 31 km per hour in densely populated suburban districts and 33 km per hour in rural areas.

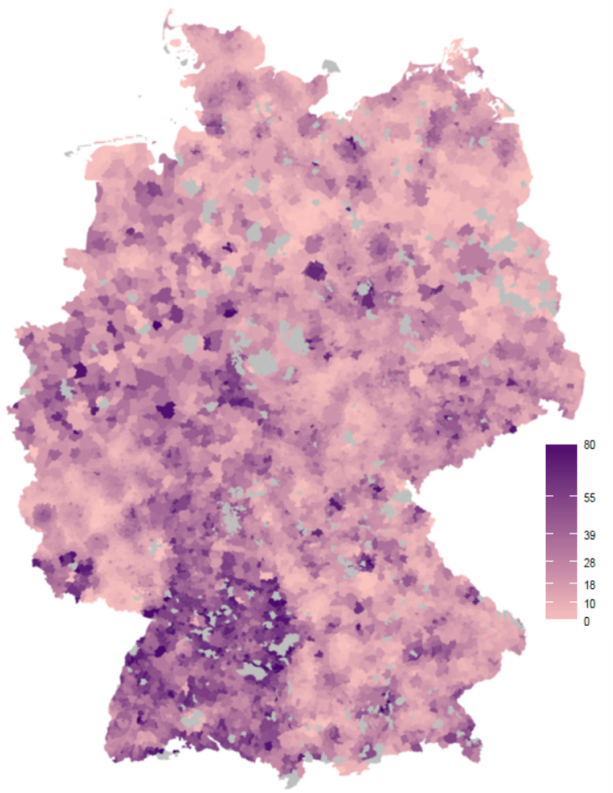
[17] In that regard, we calculate centroids (geometric centers) of all municipalities. Neighboring municipalities whose centroids are located within a given radius are taken into account in the sports expenditure measure. For the calculation of municipality centroids and the identification of municipalities within a certain radius, we used shp-files of German administrative boundaries for the years 2002-2006 provided by ©GeoBasis-DE / BKG 2017. Note that due to municipality mergers, this procedure was carried out for each year separately.

km. $w_i^{'}$ is the weight a neigboring municipality gets and the weight w_i is inverse to the distance (d_i) of the neighboring municipality to the centroid of municipality c. Weights are normalized such that they add up to 1:

$$w_i^{'} = \frac{w_i}{\sum_{j=1}^{N} w_j} \quad \text{where} \quad w_i = 1/d_i \tag{2}$$

Figure 1 displays the distance weighted expenditures on sports facilities (averaged over the years 2002-2006) for German municipalities.

Figure 1: Distance weighted per capita sports expenditures



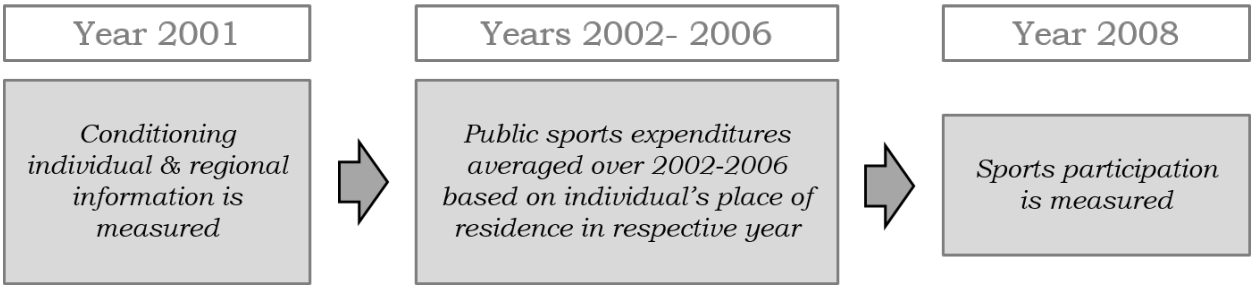
Sources: RDC of the Federal Statistical Office and Statistical Offices of the Länder (2018a; 2018b), own calculations. *Note:* Municipalities with negative expenditures or expenditures higher than the 99th-percentile are colored in gray. This figure is also contained in Pawlowski et al. (2019).

We restrict the representation of local sports expenditures to positive values up to the 99th-percentile (due to outliers). It is worth pointing out that net expenditures are comparably high in the state of Baden-Württemberg (in the south-west of Germany) and that there is also considerable within-state variation of local sports expenditures.

2.2 Individual level data: The German Socio-Economic Panel

We merge the above described regional data with individual level data of the German Socio-Economic Panel (SOEP). The SOEP is a longitudinal panel survey containing a representative sample of the German population which has been conducted yearly since 1984 (see e.g. Wagner et al., 2007).[18] We link the data sets by the official municipality key *("Amtlicher Gemeindeschlüssel" (AGS))* on an annual basis.[19] Figure 2 illustrates how we make use of this panel dataset in the analysis.

Figure 2: Timeline of measuring information



We average sports expenditures over the years 2002-2006 for each individual in order to avoid that potential year specific fluctuations in public expenditures are picked up. Individual

[18]Since SOEP data at the municipality level can only be used in the Research data center at DIW Berlin, all analyses had to be conducted there.

[19]Hence, if an individual lives in Munich from 2002-2004 and from 2005-2006 in Berlin, we assign him expenditure data from Munich for the years 2002-2004 and for the other two years from Berlin.

and regional information from the year 2001 are used as conditioning information, which we will discuss in more detail in Section 3.2.

While the SOEP provides extensive information on socio-economic characteristics, information on physical activity is not very detailed. We measure sports participation in 2008 (hence two years after the last information on sports expenditures). We do so because the survey of 2008 provides five rather than four different categories (as it is the case in 2007) with regards to frequency of sports participation. The categories in 2008 are *"daily"*, *"at least once per week"*, *"at least once per month"*, *"less than once per month"* and *"never"*. Since we want to know if expenditures on sports facilities impact the probability to practise sports, we construct a dummy variable *"sports at least monthly"* and study this outcome variable in the main results.[20]

Since we measure conditional individual and regional information in 2001 and sports participation in 2008, we restrict the sample to individuals who have a valid interview in 2001 and 2008 and are between 17 and 65 years of age in 2001 (they are in consequence between 24 and 72 years of age in 2008). We exclude individuals from the three federal city states Bremen, Hamburg and Berlin from the analysis, since there is no variation within the states composed of a single city.[21] As sports expenditures might impact sports participation differently across age groups, we estimate the effect of public sports expenditures for different age groups. Our main specification contains individuals between 31 and 45 years of age in 2001, since they have already entered the labour market in 2001 and the transition to retirement is additionally still far away when outcomes are measured in 2008. Estimation results for the younger sample (aged 17-30 in 2001) and an older sample (aged 46-65 in

[20] We also present results for the original variable in a robustness check. The results remain unchanged.
[21] All steps of constructing the SOEP sample are described in Appendix B.

2001) are discussed in Section 4.3. Our final sample contains 3,719 individuals.[22] 28% of individuals in our sample indicate that they never do sports, 23% less than monthly, 8% at least once per month, 35% at least once a week and 6% are physically active every day. Hence, 49% of the individuals in our sample participate in sports at least monthly.[23]

2.3 Descriptive statistics

Table 1 displays yearly net expenditures on sports facilities aggregated over all municipalities in which SOEP-respondents aged 17 to 65 live. Net facility expenditures amount to around 1.6 Billion per year. SOEP-respondents live in about 1,800 different German municipalities (exact numbers depend on the particular year). The total number of inhabitants in "SOEP-municipalities" account for 67% to 70% of the overall population in Germany in the respective year.[24] Average per capita facility expenditures in a municipality vary between between 26.7 and 29.2 EUR across years. Figure 3 displays a histogram of sports facility expenses in the main sample. Positive expenditures up the 95th percentile are colored in green. It is apparent that the sports expenditure variable exhibits outliers at the upper end of the distribution. Since we do not want our results to be driven by these observations, we restrict our sample to observations up to the 95th percentile and discard observations with negative or zero net expenditures. Thus, the sample is restricted to the range of values colored in green.

[22]The younger sample contains 2,015 individuals and 4,007 individuals are in the older sample.

[23]Note that the presented sports participation numbers still include individuals discarded in the last two steps of sample restriction before estimation (described in Appendix B), which encompass restrictions on the sports expenditure variable and individuals pursuing education. In the Eurobarometer Survey, 9% of respondents from Germany indicate that they exercise/play sports regularly, 40% with some regularity, 20% seldom and 31% never (European Commission, 2010). While frequency of sports participation is measured differently in both surveys, numbers from both surveys appear to be in line if one considers regular sports participation/ with some regularity as roughly equivalent to at least monthly sports participation.

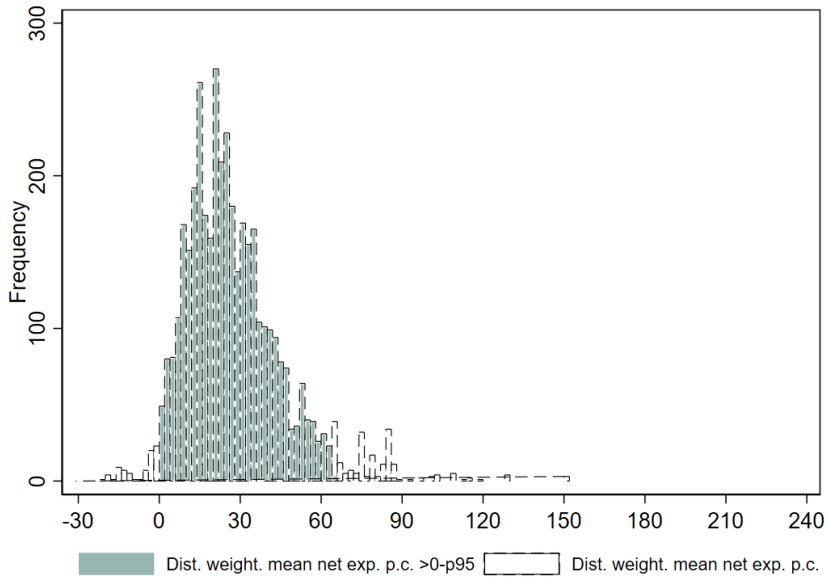
[24]Numbers are calculated based on annual population data provided by Statistisches Bundesamt (Destatis) (2018).

Table 1: Aggregated community expenditures of municipalities in SOEP sample by year

	Total expenditures facilities	Total net-expenditures facilities	Mean p.c. exp. facilities	Mean d.w.-exp. facilities	Sum of inhabitants in SOEP communities	Number of communities in SOEP sample
2002	3,297,201,875 €	1,649,571,608 €	27.1	27.4	55,281,658	1,780
2003	3,223,266,460 €	1,668,786,383 €	29.2	28.6	56,571,939	1,840
2004	3,345,680,802 €	1,568,723,536 €	28.3	28.6	57,054,481	1,877
2005	3,004,502,102 €	1,631,972,820 €	26.8	26.9	57,323,771	1,915
2006	3,192,319,697 €	1,625,980,427 €	26.7	27.1	58,324,101	1,952

Note: "Mean d.w.-exp. facilities": Mean distance weighted per capita expenditures on sports facilities. Construction of distance weighted variable as discussed above. Unlike in analyses in the rest of the paper, we did not impose restrictions on the range of values of sports expenditures for numbers presented in this table. Note that our data covers all sports expenditures channeled through municipality budgets. Hence, we do not observe expenditures directly spent by states. *Sources:* RDC of the Federal Statistical Office and Statistical Offices of the Länder (2018a; 2018b) and SOEP, own calculations.

Figure 3: Sports expenditures full range vs. restricted to positive values up to the 95th percentile



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

To get an idea how people select into communities with different sports related public expenditures, we discretise the sports expenditure variable in three states *"low"*, *"medium"*, *"high"* by cutting the variable at the 33th and 67th percentile. We then compare several regional as well as individual characteristics which we will also use later on in the analysis for these three groups in Table 2.[25]

As can be seen, compared to *"medium"* or *"low"*, *"high"* expenditures go along with lower unemployment, higher population density and less outward commuting. 28.3% of individuals in the highest tercile of sports expenditures have an upper secondary schooling degree, while this is the case for 20.5% of individuals in the lowest tercile. The share of individuals living in the southern states (Bavaria, Baden-Wuerttemberg) increases from 19.8% to 37.8% when going from the low to high sports expenditures, while the share of individuals living in the Eastern states (new federal states of Germany) decreases from 32.9% to 16.0%. Thus, it appears that higher expenses on sports facilities correlate with more prosperous and more urban areas. 48.7% of individuals in the low treatment state indicate that they do sports at least monthly compared to 54.8% in the high treatment state, hence a difference of 6 percentage points. In the rest of the paper we will try to understand how much of these differences are really affected by different sports expenditures, opposed to just go along with other features of these regions that attract active people.

[25]Table 2 contains a selection of variables used in the analysis. Descriptive statistics for the full set of variables are presented in Table B.1.

Table 2: Descriptive Statistics (selected individual and regional characteristics)

	main sample (31-45 years of age)			
	"low" expenditures	"medium" expenditures	"high" expenditures	Total
ln(population per km2/100)	1.027	1.355	1.806	1.398
	(1.275)	(1.133)	(1.074)	(1.206)
ln(average rain/100)	2.032	2.036	2.120	2.063
	(0.214)	(0.213)	(0.220)	(0.219)
ln(proportion recreational area pc in sqm/100)	-0.693	-0.856	-1.124	-0.892
	(0.766)	(0.727)	(0.670)	(0.743)
Unemployment mean 2001- 2006 (in %)	9.201	8.533	7.014	8.244
	(4.347)	(3.903)	(3.257)	(3.964)
Mean commuter flow per 1000 social sec.	-53.79	-41.33	-16.53	-37.12
employees at the workplace 2001-2006	(98.58)	(93.52)	(61.27)	(87.33)
Distance weighted sports facility	9.505	26.40	43.25	26.49
expenditures 2002 p.c. in 2004 EUR	(17.05)	(12.92)	(17.98)	(21.18)
German nationality (in %)	96.4	94.9	90.7	94.0
Upper secondary school (in %)	20.5	23.7	28.3	24.2
Northern states (in %)	17.8	13.5	3.5	11.6
Western/central states (in %)	10.7	14.1	23.7	16.2
Southern states (in %)	19.8	17.1	37.8	24.9
Eastern states (in %)	32.9	29.6	16.0	26.1
North Rhine-Westphalia (R) (in %)	18.9	25.7	19.0	21.2
Sports at least monthly 2001 (in %)	36.3	37.5	40.6	38.1
Sports at least monthly 2008 (in %)	48.7	49.9	54.8	51.1
Sports participation 2008	2.687	2.728	2.826	2.747
(1-5; 1= never, 5 = daily)	(1.343)	(1.362)	(1.360)	(1.356)

Note: Table displays sample means, standard deviations are in parenthesis. *"R"* indicates reference category in estimation. *"Northern states"*: Schleswig-Holstein, Lower Saxony; *"West/central states"*: Hesse, Saarland, Rhineland-Palatinate; *"Eastern states"*: Mecklenburg-Vorpommern, Brandenburg, Saxony-Anhalt, Saxony, Thuringia; *"Southern states"*: Baden-Wuerttemberg, Bavaria. *Sources:* RDC of the Federal Statistical Office and Statistical Offices of the Länder (2018a; 2018b), SOEP as well as further data sources as discussed in Chapter 2, own calculations.

3 Econometrics

In the following section, we outline the causal framework in the case of a continuous treatment. We then discuss the empirical strategy and its implementation given the data we have at hand. The last part of the section presents the estimation procedure.

3.1 The causal framework for continuous treatment

Let the individuals in our sample be denoted by $i = 1, ..., N$. For each observation i, a set of potential outcomes $Y_i(t)$ is defined with t ϵ $\mathcal{T}$.[26] In contrast to the binary treatment case where $\mathcal{T} = \{0, 1\}$, $\mathcal{T}$ is now an interval. Hirano and Imbens (2004) denote $Y_i(t)$ with t ϵ $\mathcal{T}$ as unit-level dose-response function and we ultimately want to estimate the average dose-response function $\mu(t) = E[Y_i(t)]$. For each unit of observation i, we observe the following: the treatment level actually received T_i, pre-treatment variables X_i assumed that they are not influenced by potential future treatment, and the outcome which corresponds to the obtained treatment level $Y_i = Y_i(T_i)$.

To estimate the dose-response function (DRF), the key assumption for identification is that selection into different levels of treatment is weakly unconfounded given X_i (introduced by Hirano and Imbens (2004) for the continuous treatment case):[27]

$$Y_i(t) \perp T_i | X_i \text{ for all t } \epsilon \mathcal{T} \tag{3}$$

Hence, conditional on observed pre-treatment variables, the level of treatment T_i is inde-

[26] This section closely follows Hirano and Imbens (2004).

[27] We also assume that a treatment assigned to an individual does not influence the outcome of another individual (Stable Unit Treatment Value Assumption) (Rosenbaum & Rubin, 1983, 1984). Additionally, we need to assume that each individual has a positive probability of obtaining each level of treatment.

pendent of the potential outcome $Y_i(t)$, which is an extension of the conditional independence assumption (CIA) in the binary-treatment case (see e.g. Heckman, LaLonde, and Smith (1999) or Imbens (2004)). Rosenbaum and Rubin (1983) show that under unconfoundedness it is sufficient to condition on the propensity score (instead of conditioning on all covariates) in the binary treatment case. Hirano and Imbens (2004) provide an extension of propensity score methods to the continuous treatment case and introduce the generalized propensity score (GPS).

The conditional density of treatment given covariates is denoted by:

$$r(t,x) = f_{T|X}\ (t|X=x) \tag{4}$$

and the GPS is the random variable $R_i = r(T_i, X_i)$, hence the conditional density at the observed treatment level. Let $R_i^t = r(t, X_i)$ be a family of random variables. The GPS is - as the propensity score - a balancing score. This means, the probability that $T = t$ is independent of X within strata with identical values of $r(t, X)$. Unconfoundedness and the balancing property of the GPS together imply that treatment assignment is unconfounded given the GPS. Hence, the GPS is able to remove biases associated with differences in covariates.

Hirano and Imbens (2004) show that if the assumption of weak unconfoundedness is fulfilled, the dose-response function can be identified as follows:

$$\begin{gathered}(i)\ \beta(t,r) = E[Y_i(t)|R_i^t = r] = E[Y_i|T_i = t, R_i = r] \\ (ii)\ \mu(t) = E[\beta(t, R_i^t)]\end{gathered} \tag{5}$$

3.2 Empirical strategy

The goal of our analysis is to investigate whether public sports expenditures affect physical activity. To answer this question, we apply a selection-on-observables approach which controls for observed confounding variables. The key assumption of the causal framework outlined in the previous section is that after conditioning on a set of covariates, selection into the level of treatment is conditionally random (Hirano & Imbens, 2004). In this section, we discuss how we aim at achieving identification given the data we have at hand.

The effect of public sports expenditures on sports participation is potentially confounded by individual characteristics as well as characteristics of the municipalities. Exemplarily, individuals most likely do not choose their place of residence randomly. Also, expenditure levels are not set at random by municipalities. In order to address this, we exploit regional information available in our data base and individual information from the SOEP. Importantly, we can only use covariates which are not affected by expenditures on sports facilities (see e.g. discussion in Lechner, 2008). We hypothesize that the following broader aspects are important when thinking about what aspects might confound the relationship of interest: First of all, economic prosperity of the community in general. Prosperous communities naturally have more scope to invest in sports infrastructure than their poorer counterparts. We account for this by controlling for unemployment averaged over the years 2001–2006 as well as the commuter flow averaged over the same time period.[28] At the same time, areas with favourable local labor market conditions are attractive for well educated people which in turn correlates with a higher likelihood of engaging in physical activity (see e.g. Humphreys and Ruseski (2015) or Lechner (2009)). As in many other countries, economic performance varies also at a higher aggregated level across regions in Germany. Exemplar-

[28] We take the average over the six years time-frame to avoid that we pick up year-specific fluctuations.

ily, while GDP per inhabitant at current prices was around 33,000–34,000 EUR in the two southern states Bavaria and Baden-Wuerttemberg in 2006, none of the new federal states of Germany had a higher GDP per inhabitant than around 21,000 EUR in 2006 (Regional Accounts VGRdL, 2018). We therefore include region dummies. *"South"* includes the states Bavaria and Baden-Wuerttemberg, *"Nord"* Lower-Saxony and Schleswig-Holstein, *"West"* Hesse, Rhineland-Palatine, Saarland and *"East"* Saxony, Saxony-Anhalt, Brandenburg and Thuringia. North Rhine-Westphalia is the reference category. We also control for population density, since investment and maintenance costs are on the one hand likely to be different in rural compared to more urban areas and urban areas are also likely to attract a different set of people.[29] Investment and maintenance costs of sports facilities are also likely to be impacted by prevailing weather conditions, which is why we also include average precipitation over the years 1981-2010. Since the recreational value of a community might influence individual location choice (it seems likely that this could be correlated with individual activity preferences) and might impact spending decisions on sports facilities of municipalities at the same time, we control for the proportion of recreational area per capita. One challenge with regards to regional characteristics arises from data availability. As mentioned in the data section, for some municipalities INKAR data is not available on the municipality level. Also, weather data is not municipality specific. Hence, some regional variables do not vary on the community level which compounds balancing (discussed in the next section). In terms of individual level characteristics, we control for a fairly standard set of socio-economic characteristics, i.e. gender, age, marital status, nationality, number of children, schooling degree and professional qualification (all information is taken from the year 2001).

As a robustness check to our baseline strategy, we exploit the panel dimension of the

[29]The role of population density on local public spending is discussed for instance by Ladd (1992) or Holcombe and Williams (2009).

data set and condition on past outcome (sports activity in 2001) as well as treatment (here we take sports expenditures from 2002) in addition to the confounding variables outlined above. In case there exist unobservable variables which display a time constant influence on sports facility expenditures and sports participation, we are able to capture them by past expenditures and past sports participation. Thereby, we in essence control for unobserved fixed effects (Lechner, 2009; Lechner & Sari, 2015) in this robustness exercise.

3.3 Estimation of the GPS

We apply the two-step estimation procedure proposed by Flores et al. (2012) to estimate the dose-response function semiparametrically.[30] The first part concerns the parametric estimation of the GPS, which is a density. We estimate the conditional density of public sports expenditures (the treatment) given covariates parametrically with Generalized Linear Models (GLM). GLM differ from "standard" linear regression models in two dimensions. First, the distribution of the dependent variable stems from the exponential family $f_y(y, \theta, \phi) = exp\{\frac{y\theta - b(\theta)}{a(\phi)} + c(y, \phi)\}$ where ϕ is the scale parameter and θ is the canonical location parameter.[31] Secondly, the GLM-framework allows for some flexibility how the conditional mean of the dependent variable is related to the explanatory variables through its "link-function" $g()$, $g\{E(Y)\} = X\gamma$, where a transformation of the mean of the dependent variable exhibits a linear relationship to the covariates and γ is a vector of coefficients (see e.g. Guardabascio & Ventura, 2014; McCullagh & Nelder, 1989). From the many possible

[30] For estimation, we use the STATA-package *"drf"* provided by Bia et al. (2014) and code by Flores et al. (2012) which we adapted at several points in order to carry out balancing as proposed in Section 3.4 and to re-estimate the GPS a second time etc.

[31] The exponential family includes, among others, the Gaussian, inverse Gaussian, Poisson, binomial, geometric, negative binomial as well as gamma distribution. For an in-depth coverage of Generalized Linear Models see e.g. Hardin and Hilbe (2012).

combinations the GLM-framework provides, we choose a specification with a Gamma distribution and a log-link, since it leads to a comparably good balancing of the covariates in our case (see results in Section 4.1). Parameter estimates are obtained via maximum likelihood. The log-likelihood function in the case of the Gamma-log model is as follows:[32]

$$\mathcal{L} = \sum_{i=1}^{n} \Big\{ \frac{T_i / exp(X_i\gamma) + (X_i\gamma)}{-\phi} + \frac{\phi + 1}{\phi} ln T_i - \frac{ln\phi}{\phi} - ln\Gamma\Big(\frac{1}{\phi}\Big) \Big\} \tag{6}$$

The estimated GPS is:[33]

$$\hat{R}_i = \frac{1}{\hat{b}_i^{\hat{a}} \Gamma(\hat{a})} T_i^{\hat{a}-1} e^{-(T_i/\hat{b}_i)} \tag{7}$$

where $\hat{b}_i = \hat{T}_i/\hat{a}$ and $\hat{a} = 1/\hat{\phi}$. As can be seen, X_i enters the estimated GPS via $E(T|X) = exp(X\gamma)$ evaluated at the estimated value of γ.[34]

3.4 Common support and balancing tests

As we use the GPS to compare people with different levels of sports expenditures but GPS of equal value, it is crucial to ensure that for each observation comparable observations with respect to the GPS can be found. We restrict the sample following the procedure proposed by Flores et al. (2012). We split the range of the sports expenditure variable into five blocks (quintiles). Each individual belongs to one of the five blocks $Q_i = (1, 2, 3, 4, 5)$. We compute

[32]Note that in the "standard" GLM-framework, which is a single parameter model, ϕ is not directly estimated. Per default, ϕ is calculated from the dispersion statistic in our case (Hardin & Hilbe, 2012; Ruoyan, 2004).

[33]For more details on GPS-estimation with an GLM-approach see e.g. Guardabascio and Ventura (2014).

[34]Note that we display the estimated GPS in the a, b-parametrization, because it is the form the STATA built-in-function *"gammaden"* computes the probability density function of the gamma distribution.

the GPS at the midpoints of the five quintiles, hence at the 10th percentile (p10), p30, p50, p70 and p90 ($\hat{R}_i^q$). Independent of to which block an individual belongs to, we obtain five evaluations of the GPS for each observation. The common support condition for block q is as follows:

$$CS_q = \Big\{ i : \hat{R}_i^q \; \epsilon \; \Big[max\Big\{ min_{j:Q_j=q} \hat{R}_j^q, min_{j:Q_j \neq q} \hat{R}_j^q \Big\}, \\ min\Big\{ max_{j:Q_j=q} \hat{R}_j^q, max_{j:Q_j \neq q} \hat{R}_j^q \Big\} \Big] \Big\} \tag{8}$$

This means, for each of the five blocks, we make a comparison between the support of the distribution for individuals in the respective block ($Q_i = q$) and individuals in all other blocks ($Q_i \neq q$). Finally, we only keep individuals who fulfil CS_1 through CS_5.

Additionally, it is necessary to evaluate whether the estimated GPS balances the covariates well. In this section, we suggest an approach to assess balancing in the continuous treatment case which differs from the ones previously applied in the literature. Beforehand, we will briefly present the core methods used so far.

Kluve et al. (2012) draw comparison between significance of coefficients regressing each covariate on the treatment variable with and without conditioning on the distribution of the GPS, whereby the GPS is evaluated at different percentiles of the distribution of the treatment variable (they choose the 25th, the 50th as well as the 75th percentile). Their line of reasoning is that after conditioning on the distribution of the GPS, the treatment should be not correlated with the covariate anymore. Flores et al. (2012) estimate an unrestricted model for T_i which contains the GPS up to a cubic term as well as covariates. In the second model ("restricted model") all covariates are set to zero. They use a likelihood ratio test to compare the unrestricted with the restricted model.

Hirano and Imbens (2004) discretize the treatment variable in several groups and predict the GPS for each individual at the median value within each group. Each GPS-prediction is divided into five quantiles. In a second step, they compare means of observations who find themselves in the same quantile of GPS-prediction, but are not in the same treatment group. They calculate weighted averages out of these five differences in means and do t-tests on the sample means on the unadjusted as well as on the sample means adjusted for the GPS. One problem of t-tests or other statistical tests is their sensitivity to sample size (see for example the discussion in Austin (2009) for the binary treatment case). Austin (2018) recently proposed to alternate the last two steps of the approach by Hirano and Imbens (2004) to calculate standardized differences instead of doing t-tests on the sample means.

While we also suggest to assess balancing via standardized differences, we apply a different aggregation procedure, introduce an additional common support step and do pairwise comparisons of all groups instead of comparing one group against all other groups collectively. Pairwise comparison of all groups is important, because lumping the groups together might obscure imbalances. This was also noted by Austin (2018) as a limitation of the approach by Hirano and Imbens (2004).

Rosenbaum and Rubin (1985) introduced the concept of standardized differences for assessing balancing in propensity score matching for the binary treatment case. In the binary treatment case, standardized difference before and after matching are calculated as follows (1 treated, 0 untreated):

$$SD_{before\ Binary}(X) = 100 * \frac{\bar{X}_1 - \bar{X}_0}{\sqrt{\left[V_1(X) + V_0(X)\right]/2}} \tag{9}$$

$$SD_{after\ Binary}(X) = 100 * \frac{\bar{X}_{1,M} - \bar{X}_{0,M}}{\sqrt{\left[V_1(X) + V_0(X)\right]/2}} \tag{10}$$

As can be seen, standardized differences after matching is the difference in means in the common support sample divided by the square root of the mean sample variances of the full sample.

In the continuous treatment case, assessing balancing via standardized differences is considerably more complex. With regards to standardized differences before GPS adjustment, we make pairwise comparisons between the 5 blocks, hence Block 1 versus Block 2 etc. and consequently get 5*4/2 = 10 comparisons. Standardized differences are calculated as follows:[35]

$$SD_{before}(X)^{q,q'} = 100 * \frac{\bar{X}^q - \bar{X}^{q'}}{\sqrt{\left[V(X^q) + V(X^{q'})\right]/2}} \tag{11}$$

With respect to standardized differences after GPS-adjustment, we start by splitting each of the five GPS predictions into 5 quantiles.[36] There exist 5^5 possible combinations of cells for an observation in terms of the 5 GPS evaluations at p10, p30, p50, p70 and p90. In our case, 21 cells out of 3,125 possible cells contain observations. We keep individuals in cells c where observations from all five blocks (hence the whole range of sports expenditures) are represented and in doing so, we introduce an additional common support condition. For each quintile q, we calculate the mean in each cell and aggregate over all cells. The means are weighted by the number of observations in the respective cell (N_c) over all observations of the common support sample (N_s). As in the unweighted case, we have 5*4/2 comparisons

[35]Note that we report mean differences for binary variables.

[36]This first step is hence identical to Hirano and Imbens (2004).

to make:

$$SD_{after}(X)^{q,q'} = 100 * \frac{\Big[\sum_{c=1}^{C}(\frac{N_c}{N_s}\bar{X}_s^q)\Big] - \Big[\sum_{c=1}^{C}(\frac{N_c}{N_s}\bar{X}_s^{q'})\Big]}{\sqrt{\big[V(X^q) + V(X^{q'})\big]/2}} \tag{12}$$

For binary variables, we calculate differences of weighted means. After this procedure, we re-estimate the GPS on the common support sample and impose the overlap condition a second time.[37]

Results of the employed common support procedure are discussed in Section 4.1. Table 3 summarizes the steps of the proposed common support procedure and indicates the number of observations discarded in each step.

Table 3: Common support procedure

	Steps of common support procedure	**Numbers of observations discarded (Number of observations left)**
Step 1	Estimation of GPS and imposition of common support condition	-126 observations (N = 3,593)
Step 2	Keep cells which contain observations from all five blocks	-23 observations (N= 3,570)
Step 3	Re-estimation of GPS on common support sample and imposition of common support condition for the second time	-38 observations (N = 3,532)

Sources: RDC of the Federal Statistical Office and Statistical Offices of the Länder (2018a; 2018b), SOEP as well as further data sources as discussed in Chapter 2, own calculations.

Assessment of balancing in this way is appealing for several reasons: First of all, stan-

[37]Please note that we do not carry out the balancing step a second time in which additional observations were discarded in the first round.

dardized differences are in contrast to t-tests or other statistical tests not sensitive to sample size. In addition, by deleting cells where not all blocks are represented, the procedure acts as an additional common support condition. Finally, pairwise comparisons of all blocks instead of comparing one block to all other blocks serves as an useful diagnostic tool, since it helps to identify where problems of balancing are arising from.

3.5 Estimation of the dose-response function

Finally, the dose-response function $\mu(t)$ is estimated via local linear regression of sports participation (Y) on sports expenditures (T) with a weighted kernel function defined as: $\tilde{K}_{h,X}(T_i - t) = K_h(T_i - t)/\hat{R}_i^t$. $K(u)$ is a kernel function and h is a bandwidth that asymptotically satisfies $h \to 0$ and $Nh \to \infty$ as $N \to \infty$, $K_h(z) = h^{-1}K(z/h)$.[38] As can be seen, the kernel weight is not divided by an estimate of an individual's "own GPS" $\hat{R}_i$, but rather by the individual-specific estimate of her GPS at a point t of the distribution of T. The estimator of the average dose-response function is:

$$\hat{\mu}(t) = \frac{D_0(t)S_2(t) - D_1(t)S_1(t)}{S_0(t)S_2(t) - S_1^2(t)} \tag{13}$$

with $S_j(t) = \sum_{i=1}^{N} \tilde{K}_{h,X}(T_i - t)(T_i - t)^j$ and $D_j(t) = \sum_{i=1}^{N} \tilde{K}_{h,X}(T_i - t)(T_i - t)^j Y_i$, j = 0,1,2[39]

Besides the dose-response function, we estimate the treatment effect function $\mu(t + x) - \mu(t)$. The estimates for the dose-response function are obtained at 99 distinct values of local

[38]For all details of the estimation procedure see Flores et al. (2012).

[39]The estimator uses a normal kernel and the global bandwidth is chosen by applying the approach suggested by Fan and Gijbels (1996). Their procedure suggests to take a global polynomial of order p+3 (p denotes the order of the fitted local polynomial) for the estimation of the unknown terms in the optimal global bandwidth.

expenditures on sports facilities. The chosen values match the percentiles of the empirical distribution of the sports expenditure variable. We cluster on municipalities in the pre-treatment period (2001) and calculate standard errors with bootstrap (1,000 replications).[40]

4 Results

In this section, we first discuss results of the estimation of the GPS - hence the first step of the estimation procedure. We then present a graphical representation of the employed common support procedure as well as results on the balancing tests. Finally, we present results on the impact of sports expenditures on sports participation and several robustness checks.

4.1 Estimation of the GPS, common support & balancing tests

Results of the GPS estimation exhibit a similar pattern as the descriptive statistics displayed in Table 2: Higher population density, less unemployment and less outward commuting are associated with higher sports facility expenditures (results are displayed in Table C.1). With the exception of the proportion of recreational space, coefficients of all regional variables are significant in the estimation. With respect to individual characteristics, having an upper secondary school degree or an intermediate secondary school degree is positively correlated with sports expenditures, whereas German nationality displays a negative correlation with sports facilities expenses. This is coherent with the fact that education levels are higher in urban and economically thriving areas.

[40]Note that we do not account for potential regional correlation due to computational and time constraints at the research data center.

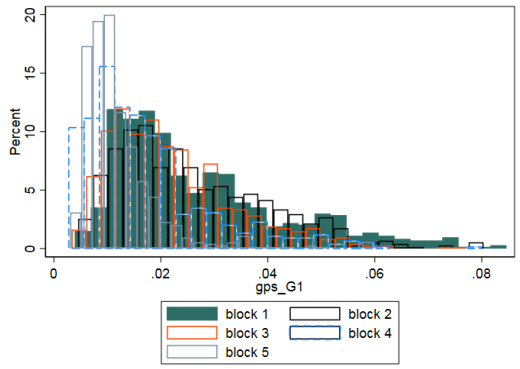
Figure 4 presents results from the employed common support procedure. While step 1 and step 3 of the procedure (summarized in Table 3) follow Flores et al. (2012), we choose a different graphical representation. Flores et al. (2012) show overlaid histograms of the "own" GPS $\hat{R}_i$ in one block versus all other blocks. We on the other hand provide histograms of the distribution of the predicted GPS for the five blocks before and after imposing the common support condition at the five points where we evaluate the GPS (p10, p30, p50, p70 and p90). Potential common support problems are hereby easier detectable and identifiable. Figure 4 displays the overlaid histograms of the five blocks at p10 and p30.[41] For both evaluation points, Figures are ordered from top to bottom (before common support, after step 1 and after step 3). It is visible from the figures that extreme values become discarded through the process.

We next discuss results from our suggested approach to assess balancing in the continuous treatment case. Table 4 contains a selection of pair-wise comparisons of blocks before GPS-adjustment. Large standardized biases can be observed especially for all regional variables.[42] Not surprisingly, these biases are particularly high when we compare blocks "far apart" (for instance Block 1 versus Block 5 or Block 2 versus Block 5). For instance, 18% of individuals in the lowest quintile live in the two southern states, whereas this is the case for 43% of individuals in the highest quintile, hence a difference of 25 percentage points (the difference is 23 percentage points when we compare Block 2 versus Block 5). Balancing results after step 2 of the common support procedure are presented in Table 5. While balancing is still not ideal, it has improved considerably compared to before GPS-adjustment. For instance, while the standardized difference for population density was -67.29 before GPS-adjustment (Block 2 versus Block 5), it is down to -5.89 after GPS-adjustment.

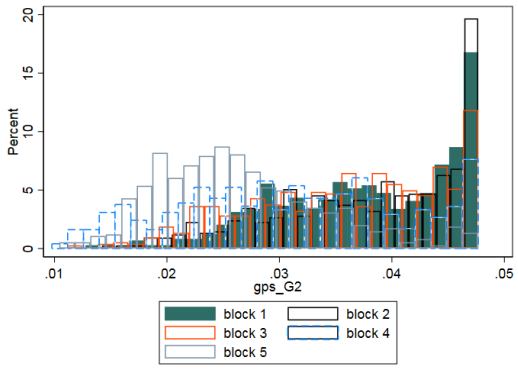
[41] Figures for the three other points of evaluation can be found in Appendix C.

[42] The full set of all pairwise-comparisons are contained in Table C.2 in Appendix C.

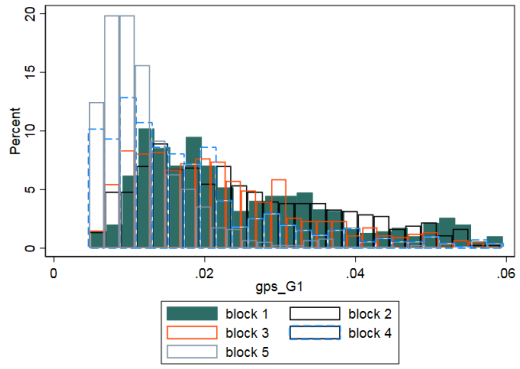
Figure 4: Common support



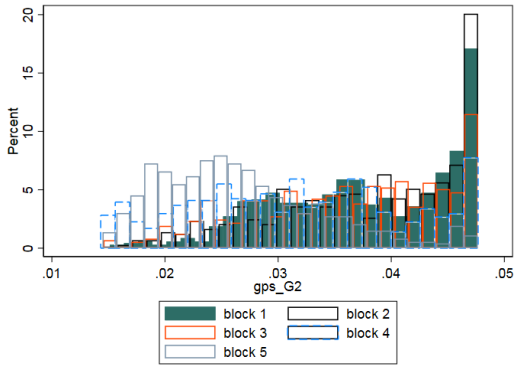
(a) GPS at p10 - before CS condition



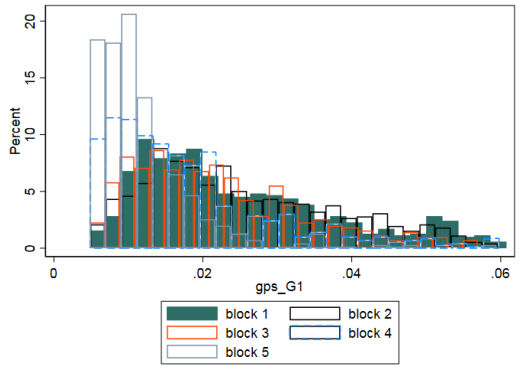
(b) GPS at p30 - before CS condition



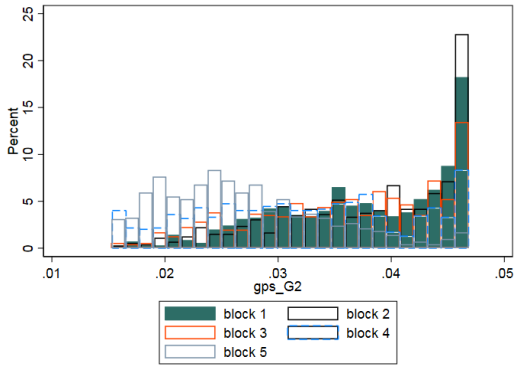
(c) GPS at p10 - after step 1



(d) GPS at p30 - after step 1



(e) GPS at p10 - after step 3



(f) GPS at p30 - after step 3

Note: Figures (e) and (f) show overlaid histograms of the blocks after re-estimation of the GPS and imposing the common support condition a second time. *Sources:* RDC of the Federal Statistical Office and Statistical Offices of the Länder (2018a; 2018b), SOEP as well as further data sources as discussed in Chapter 2, own calculations.

Table 4: Balancing before GPS adjustment

	Block 1 vs 2			Block 1 vs 5			Block 2 vs 5		
	Mean 1	Mean 2	% bias	Mean 1	Mean 5	% bias	Mean 2	Mean 5	% bias
ln(population per km2/100)	1.15	1.07	6.54	1.15	1.81	-56.21	1.07	1.81	-67.29
ln(average rain/100)	2.02	2.02	1.94	2.02	2.14	-53.97	2.02	2.14	-56.98
ln(proportion recreational area pc in sqm/100)	-0.68	-0.73	7.12	-0.68	-1.17	71.48	-0.73	-1.17	58.26
Unemployment mean	9.26	9.07	4.41	9.26	6.40	78.06	9.07	6.40	74.75
Commuters mean	-45.12	-50.32	5.60	-45.12	-16.67	-36.85	-50.32	-16.67	-41.74
Age	38.03	38.36	-7.81	38.03	37.73	6.92	38.36	37.73	14.97
Number of children	1.27	1.32	-5.05	1.27	1.23	3.27	1.32	1.23	8.46
Male	0.47	0.46	0.01	0.47	0.47	0.00	0.46	0.47	-0.01
Married	0.71	0.74	-0.03	0.71	0.73	-0.02	0.74	0.73	0.02
Divorced	0.09	0.09	0.00	0.09	0.08	0.01	0.09	0.08	0.01
German	0.95	0.96	0.00	0.95	0.87	0.08	0.96	0.87	0.09
Other graduation diploma	0.04	0.05	-0.01	0.04	0.09	-0.05	0.05	0.09	-0.04
Intermediate secondary school	0.45	0.49	-0.04	0.45	0.34	0.11	0.49	0.34	0.15
Upper secondary school	0.21	0.19	0.01	0.21	0.28	-0.08	0.19	0.28	-0.09
No vocational degree	0.11	0.10	0.02	0.11	0.16	-0.04	0.10	0.16	-0.06
Degree below university	0.79	0.80	-0.01	0.79	0.71	0.07	0.80	0.71	0.09
University	0.18	0.22	-0.04	0.18	0.22	-0.04	0.22	0.22	0.00
Northern states	0.21	0.16	0.05	0.21	0.02	0.19	0.16	0.02	0.14
Western states	0.11	0.09	0.02	0.11	0.24	-0.12	0.09	0.24	-0.15
Southern states	0.18	0.20	-0.03	0.18	0.43	-0.25	0.20	0.43	-0.23
Eastern states	0.30	0.36	-0.05	0.30	0.09	0.21	0.36	0.09	0.27

Sources: RDC of the Federal Statistical Office and Statistical Offices of the Länder (2018a; 2018b), SOEP as well as further data sources as discussed in Chapter 2, own calculations.

Table 5: Balancing after step 2 of common support procedure

	Block 1 vs 2	Block 1 vs 3	Block 1 vs 4	Block 1 vs 5	Block 2 vs 3	Block 2 vs 4	Block 2 vs 5	Block 3 vs 4	Block 3 vs 5	Block 4 vs 5
ln(population per km2/100)	1.91	-8.42	-12.22	-3.45	-10.97	-14.84	-5.89	-4.47	6.07	10.54
ln(average rain/100)	15.90	17.42	16.01	19.59	1.02	1.35	3.55	0.45	2.64	1.91
ln(Prop. recreational area pc in sqm/100)	15.12	7.02	18.16	-10.43	-8.91	0.95	-25.30	11.21	-18.03	-30.08
Unemployment mean	0.34	6.17	-4.31	2.52	5.93	-4.77	2.18	-11.29	-4.71	8.11
commuters mean	0.98	5.51	-6.01	6.07	4.38	-6.86	4.69	-12.03	-0.45	14.52
Age in years	-4.78	5.05	4.15	-3.97	10.00	9.26	0.74	-1.07	-9.09	-8.31
Number of children	-6.15	-6.98	-6.66	8.06	-1.05	-0.72	14.53	0.32	15.10	14.80
Male	-0.02	-0.03	-0.03	0.00	-0.01	-0.01	0.02	0.01	0.03	0.03
Married	0.00	-0.04	0.00	0.00	-0.03	0.00	0.00	0.03	0.03	0.00
Divorced	0.00	0.03	0.02	0.01	0.03	0.01	0.01	-0.01	-0.02	-0.01
German	0.01	0.02	0.00	0.01	0.01	-0.01	0.00	-0.02	-0.01	0.01
Other graduation diploma	-0.01	-0.01	-0.01	0.00	-0.01	0.00	0.01	0.01	0.02	0.01
Intermediate secondary school	-0.01	0.04	0.00	0.01	0.05	0.01	0.02	-0.04	-0.03	0.01
Upper secondary school	0.02	-0.03	0.02	-0.03	-0.04	0.00	-0.05	0.04	0.00	-0.05
No vocational degree	0.01	0.01	0.00	0.02	0.00	-0.01	0.01	-0.01	0.01	0.02
Degree below university	0.00	0.00	0.00	-0.01	0.01	0.01	-0.01	0.00	-0.01	-0.01
University degree	-0.04	0.00	-0.02	-0.06	0.04	0.02	-0.02	-0.02	-0.06	-0.04
Northern states	0.03	0.01	0.01	0.03	-0.03	-0.02	0.00	0.00	0.03	0.02
Western states	0.03	-0.02	0.01	0.02	-0.05	-0.01	-0.01	0.04	0.04	0.01
Southern states	-0.03	0.05	0.02	0.03	0.08	0.05	0.07	-0.03	-0.01	0.02
Eastern states	-0.06	0.00	-0.05	-0.05	0.06	0.01	0.01	-0.05	-0.05	0.00

Sources: RDC of the Federal Statistical Office and Statistical Offices of the Länder (2018a; 2018b), SOEP as well as further data sources as discussed in Chapter 2, own calculations.

4.2 Effects of sports expenditures on sports participation

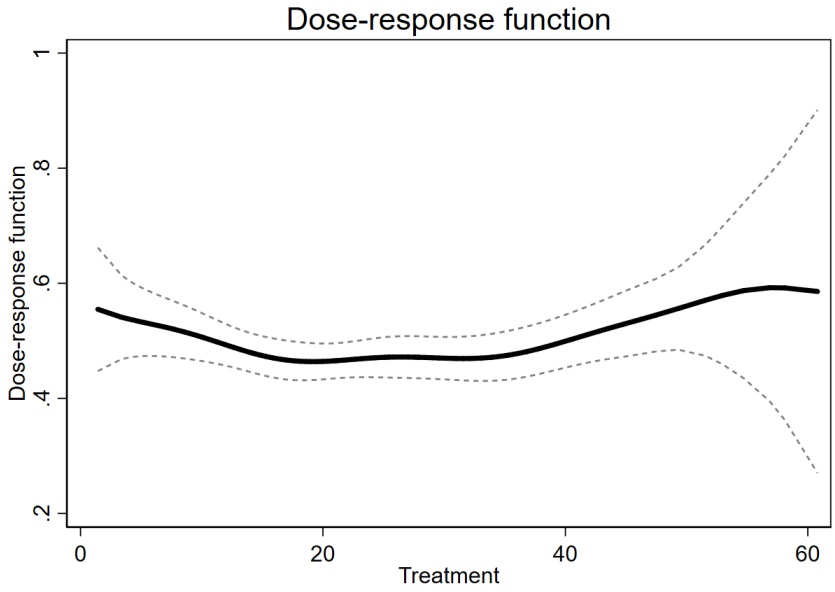
Our main result concerns the dose-response relationship of expenditures on sports facilities on the binary outcome variable *"sports at least monthly"* measured in 2008. A set of robustness exercises is discussed in the next section. Figure 5a shows the estimated dose-response function (DRF) and Figure 5b the corresponding treatment effect function[43] (both with pointwise 95% confidence bands). We find that the dose-response function is quite flat in the range of values where we are able to obtain precise estimates and sports participation is neither an in- or decreasing function of the amount of sports facility expenses as can be seen in Figure 5a. Turning to the treatment effect function, we observe no estimates that are statistically different from zero. As expected, confidence intervals become wider at the upper end of the distribution where we have few data points.

To illustrate the lacking influence of increased levels of sports expenditures, we compare estimates for the dose-response function for 20 EUR and 30 EUR per capita expenditures:[44] At percentile 37 (which corresponds to net per capita sports facility expenditures of 19.95 EUR) we obtain an estimate for the DRF of 0.464. At percentile 65 (30.1 EUR net sports facility expenditures per capita), the corresponding estimate is 0.470. Hence, increasing per capita expenditures from roughly 20 EUR to 30 EUR per capita would only lead to a change in at least monthly sports participation of 0.6 percentage points. Going from the lower bound of the confidence interval at percentile 37 (0.433) to the upper bound of the confidence interval at percentile 65 (0.507), which gives us an idea how large the effect could possibly be while staying within the confidence bands, would translate into an increase in the probability of sports participation by 7.4 percentage points.

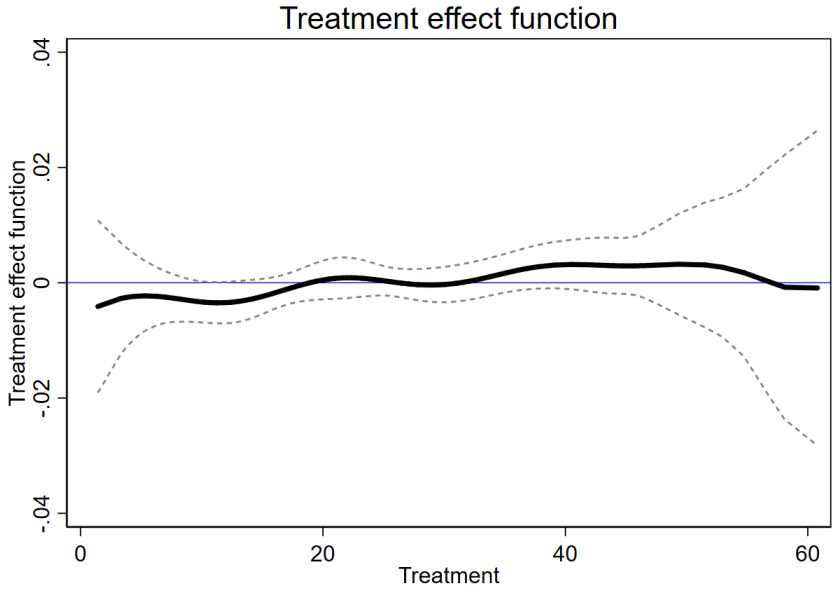
[43] Calculated as $\mu(t + 0.5) - \mu(t)$.

[44] Note that median sports facility expenditures are 23.82 EUR per capita.

Figure 5: Results main specification



(a) Dose-response function "sports at least monthly"



(b) Treatment effect function "sports at least monthly"

Sources: RDC of the Federal Statistical Office and Statistical Offices of the Länder (2018a; 2018b), SOEP as well as further data sources as discussed in Chapter 2, own calculations.

4.3 Robustness

As a first robustness check, we estimate the relationship of interest for the "raw" sports variable which entails five categories, a scaled version of the sports variable (since the distances between the different categories of the sports variable are not equal) and for a later point in time (sports participation measured in 2013). The coding of the scaled sports variable is as follows: *"never" = 0*, *"less than monthly" = 1*, *"at least once a month" = 5*, *"at least once a week" = 27*, *"daily"= 60*.[45] Irrespective of measuring the outcome variable as a scaled version, taking the "raw" version of the variable ranging from 1-5 or looking at a later point in time, we do not find an impact of sports facility expenditures on sports participation as can be seen in Figure D.1. The left panel of the Figure contains the estimated dose-response function and the right panel the corresponding treatment effect function.

Since our main specification contains individuals from 31-45 years of age, we check if sports expenditures impact sports participation in the younger (A1) and the older age group (A3) (Figures D.2 and D.3). We do not find an impact of sports facility expenditures on sports participation here neither. For the group of individuals aged 46-65 (A3), raising per capita expenditures from roughly 20 to 30 EUR per capita would only lead to a change in at least monthly sports participation (first row left panel of Figure D.3) of 0.2 percentage points (compared to 0.6 percentage points in the main specification).[46]

We include lagged sports participation and lagged facility expenditures for reasons dis-

[45] The reasoning for scaling the variable in this way is as follows: A two months period equals roughly 60 days. Hence, daily exercise is set to 60. Correspondingly, two months are approximately 9 weeks and we hence rescaled physical exercise "at least once a week" to $9 * 3 = 27$ by choosing 3 times a week for this category etc. Of course, it would be important to check the robustness of the chosen scaling. However, since the results looked very similar for the unscaled and scaled version of the outcome variable we refrained from doing so due to time and computing power constraints at the research data center.

[46] Note that for outcomes other than the binary sports outcome variable in 2008, we only present the treatment effect function in the paper. Respective dose-response functions can be found in Appendix E.

cussed in Section 3.2 in a next step. Figure D.4 shows the dose-response functions and treatment effect functions for all three age groups and the binary outcome variable *"sports at least monthly"*. We focus on comparing estimates for our main sample (individuals from 31 to 45 years of age) with and without conditioning on lagged outcome and treatment in the following. Imposing common support and the balancing condition (each cell needs to contain individuals from all five blocks) reduces our estimation sample by 1,344 observations, hence roughly 33% of the sample (compared to 149 observations in the main specification without lagged treatment and outcome). Hence, our common support sample is considerably smaller compared to the sample in the main specification. Results are presented in Figure D.4. The range of values where we get precise estimates is smaller (as expected). Estimates for the dose-response function are quite similar in size compared to estimates in the main specification though. At 19.9 EUR (percentile 43), we obtain an estimate for the DRF of 0.460. At 30.0 EUR p.c. expenditures, the estimate is 0.469. Consequently, raising per capita expenditure from 20 EUR to 30 EUR would correspond to a change in at least monthly physical activity of 0.9 percentage points.

As a final robustness check, we split our sample by municipality size. The reasoning behind this is that one could expect that spending decisions might be undertaken differently in smaller compared to very large municipalities (in terms of number of inhabitants). Median municipality size in our main specification is 21,671 inhabitants. Splitting the sample at 20,000 inhabitants we do not find any heterogeneity regarding the effect of sports facility expenditures in terms of municipality size.

5 Conclusion

We exploit a combined data set of municipality expenditures and SOEP individual level data to study the impact of local expenditures on sports facilities on the individual sports participation of adults. The resulting dose-response relationships show no influence of facility expenditures on individual sports participation. Our results are robust to different specifications of the outcome variable, measuring sports participation at a later point in time, looking at different age groups, the inclusion of lagged outcome plus treatment as well as splitting the sample into small and large municipalities. This study shows that expenditures on sports facilities were not an effective tool to push adult sports participation forward in the German context.

Our analysis faces some limitations which are at the same time important topics for future research. First of all, our measure for sports participation only provides information on the frequency of sports participation. It neither contains information on the time spent on exercising, nor the type of sports or the intensity of exercise. Hence, it might well be that higher expenditures on sports facilities lead to changes in one of these aspects not covered in the survey. Unfortunately, we cannot improve on that dimension since to the best of our knowledge the SOEP is currently the only representative sample for Germany which is available on the municipality level and that provides information on sports participation. While an analysis on the impact of expenditures on sports facilities for children is beyond the scope of this paper, it would be interesting to study this in further research and investigate whether the picture looks similar for them.

References

Austin, P. C. (2009). Balance diagnostics for comparing the distribution of baseline covariates between treatment groups in propensity-score matched samples. *Statistics in Medicine*, *28*(25), 3083–3107.

Austin, P. C. (2018). Assessing the performance of the generalized propensity score for estimating the effect of quantitative or continuous exposures on binary outcomes. *Statistics in Medicine*, *37*(11), 1874–1894.

BBSR (Bundesinstitut für Bau-, Stadt- und Raumforschung). (2013). *Indikatoren und Karten zur Raum- und Stadtentwicklung in Deutschland und Europa (CD-Rom).* Bonn: BBSR.

Bia, M., Flores, C. A., Flores-Lagunes, A., & Mattei, A. (2014). A Stata package for the application of semiparametric estimators of dose–response functions. *The Stata Journal*, *14*(3), 580–604.

Choe, C., Flores-Lagunes, A., & Lee, S.-J. (2015). Do dropouts with longer training exposure benefit from training programs? Korean evidence employing methods for continuous treatments. *Empirical Economics*, *48*(2), 849–881.

Dallmeyer, S., Wicker, P., & Breuer, C. (2017). Public Expenditure and Sport Participation: An Examination of Direct, Spillover, and Substitution Effects. *International Journal of Sport Finance*, *12*(3), 244–264.

Dammert, A. C., & Galdo, J. (2013). Program quality and treatment completion for youth training programs. *Economics Letters*, *119*(3), 243–246.

Downward, P., & Rasciute, S. (2015). Exploring the covariates of sport participation for health: an analysis of males and females in England. *Journal of Sports Sciences*, *33*(1), 67–76.

DWD (Deutscher Wetterdienst). (2016). *Langjährige Stationsmittelwerte für die Klimareferenzperiode 1981–2010 für den aktuellen Standort und Bezugsstandort [Long-year station average values for the climatic reference period 1981–2010 for the current location and the reference location]. Version V0.x.*

European Commission. (2010). Sport and physical activity. In *Special Eurobarameter 334/ Wave 72.3 - TNS Opinion & Social.* Brussels: European Commission.

Fan, J., & Gijbels, I. (1996). *Local Polynomial Modelling and Its Applications.* New York:

Chapman and Hall/CRC.

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

Federal Statistical Office. (2011). *Gemeindeverzeichnis. Gebietsstand: 31.12.2010.*

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 2nd November 2017]*. `https://www-genesis.destatis.de/genesis/online?sequenz=statistikTabellen&selectionname=61111`, accessed: 2 November 2017.

Flores, C. A., Flores-Lagunes, A., Gonzalez, A., & Neumann, T. C. (2012). Estimating the Effects of Length of Exposure to Instruction in a Training Program: The Case of Job Corps. *Review of Economics and Statistics*, *94*(1), 153–171.

German Federal Government. (2006). *FIFA World Cup 2006. Final report of the Federal Government [Fußball-WM 2006. Abschlussbericht der Bundesregierung]*. `https://www.gudrun-zollner.de/files/gudrun_zollner/berlin/Broschueren_Sport/WM2006_Abschlussbericht_der_Bundesregierung.pdf`, retrieved October 26, 2017.

Guardabascio, B., & Ventura, M. (2014). Estimating the dose–response function through a generalized linear model approach. *The Stata Journal*, *14*(1), 141–158.

Guthold, R., Stevens, G. A., Riley, L. M., & Bull, F. C. (2018). Worldwide trends in insufficient physical activity from 2001 to 2016: a pooled analysis of 358 population-based surveys with 1· 9 million participants. *The Lancet Global Health*, *6*(10), e1077–e1086.

Hardin, J. W., & Hilbe, J. M. (2012). *Generalized Linear Models and Extensions, Third Edition*. Stata Press.

Heckman, J. J., LaLonde, R. J., & Smith, J. A. (1999). The Economics and Econometrics of Active Labor Market Programs. In O. Ashenfelter & D. Card (Eds.), *Handbook of Labor Economics* (Vol. 3A, pp. 1865–2097). Elsevier.

Hirano, K., & Imbens, G. W. (2004). The propensity score with continuous treatments. *Applied Bayesian modeling and causal inference from incomplete-data perspectives*, *226164*, 73–84.

Holcombe, R. G., & Williams, D. W. (2009). Are there economies of scale in municipal government expenditures? *Public Finance and Management*, *9*(3), 416.

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

Humphreys, B. R., & Ruseski, J. E. (2015). The Economic Choice of Participation and Time Spent in Physical Activity and Sport in Canada. *International Journal of Sport Finance*, *10*(2), 138–159.

Imai, K., & Van Dyk, D. A. (2004). Causal inference with general treatment regimes: Generalizing the propensity score. *Journal of the American Statistical Association*, *99*(467), 854–866.

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

Imbens, G. W. (2004). Nonparametric estimation of average treatment effects under exogeneity: A review. *Review of Economics and Statistics*, *86*(1), 4–29.

infas, DLR. (2010). *Mobilität in Deutschland 2008 – Ergebnisbericht: Struktur, Aufkommen, Emissionen, Trends [Mobility in Germany 2008 – Result report].* Bonn/Berlin: Infas / DLR.

Kluve, J., Schneider, H., Uhlendorff, A., & Zhao, Z. (2012). Evaluating continuous training programmes by using the generalized propensity score. *Journal of the Royal Statistical Society: Series A (Statistics in Society)*, *175*(2), 587–617.

Kokolakakis, T., Lera-López, F., & Castellanos, P. (2014). Regional differences in sports participation: The case of local authorities in England. *International Journal of Sport Finance*, *9*(2), 149–171.

Ladd, H. F. (1992). Population growth, density and the costs of providing public services. *Urban Studies*, *29*(2), 273–295.

Langer, M. (2006). *Öffentliche Förderung des Sports: eine ordnungspolitische Analyse (Volkswirtschaftliche Schriften 548.* Berlin: Duncker & Humblot.

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

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

Lechner, M., & Sari, N. (2015). Labor market effects of sports and exercise: Evidence from Canadian panel data. *Labour Economics*, *35*, 1–15.

McCullagh, P., & Nelder, J. A. (1989). *Generalized Linear Models* (2nd ed., Vol. 37). Chapman and Hall, London.

Moodie, E. E., & Stephens, D. A. (2012). Estimation of dose–response functions for longitudinal data using the generalised propensity score. *Statistical Methods in Medical Research*, *21*(2), 149–166.

Pawlowski, T., & Breuer, C. (2012). *Die finanzpolitische Bedeutung des Sports in Deutschland*. Wiesbaden: Springer-Gabler Research.

Pawlowski, T., Breuer, C., Wicker, P., & Poupaux, S. (2009). Travel time spending behaviour in recreational sports: an econometric approach with management implications. *European Sport Management Quarterly*, *9*(3), 215–242.

Pawlowski, T., Steckenleiter, C., Wallrafen, T., & Lechner, M. (2019). *Individual labor market effects of local public expenditures on sports.* (mimeo)

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

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

Regional Accounts VGRdL. (2018). *Gross domestic product at current prices, 1991 to 2017 (WZ2008).* `https://www.statistik-bw.de/VGRdL/tbls/tab.jsp?lang=en-GB&rev=RV2014&tbl=tab01`, accessed: 4 January 2019.

Reimers, A. K., Wagner, M., Alvanides, S., Steinmayr, A., Reiner, M., Schmidt, S., & Woll, A. (2014). Proximity to sports facilities and sports participation for adolescents in Germany. *PLoS One*, *9*(3), e93059.

Rosenbaum, P. R., & Rubin, D. B. (1983). The central role of the propensity score in observational studies for causal effects. *Biometrika*, *70*(1), 41–55.

Rosenbaum, P. R., & Rubin, D. B. (1984). Reducing bias in observational studies using sub-

classification on the propensity score. *Journal of the American Statistical Association*, *79*(387), 516–524.

Rosenbaum, P. R., & Rubin, D. B. (1985). Constructing a control group using multivariate matched sampling methods that incorporate the propensity score. *The American Statistician*, *39*(1), 33–38.

Ruoyan, M. (2004). *Estimation of dispersion parameters in GLMs with and without random effects*. Citeseer.

Statistisches Bundesamt (Destatis). (2018). *Population.* `https://www.destatis.de/DE/ZahlenFakten/GesellschaftStaat/Bevoelkerung/Bevoelkerungsstand/Tabellen_/lrbev01.html`, accessed: 21 August 2018.

Steinmayr, A., Felfe, C., & Lechner, M. (2011). The closer the sportier? Children's sports activity and their distance to sports facilities. *European Review of Aging and Physical Activity*, *8*(2), 67-82.

Tiebout, C. M. (1956). A pure theory of local expenditures. *Journal of Political Economy*, *64*(5), 416–424.

Wagner, G. G., Frick, J. R., & Schupp, J. (2007). The German Socio-Economic Panel Study (SOEP)–Scope, Evolution and Enhancements. *Schmollers Jahrbuch: Journal of Applied Social Science Studies/Zeitschrift für Wirtschafts-und Sozialwissenschaften*, *127*(1), 139–169.

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

World Health Organization. (2013). Global action plan for the prevention and control of noncommunicable diseases 2013-2020 [internet]. *Geneva: WHO*.

World Health Organization. (2018). *Noncommunicable diseases.* `http://www.who.int/news-room/fact-sheets/detail/noncommunicable-diseases`, accessed: 21 August 2018.

Wu, X., Mealli, F., Kioumourtzoglou, M.-A., Dominici, F., & Braun, D. (2018). Matching on Generalized Propensity Scores with Continuous Exposures. *arXiv preprint arXiv:1812.06575*.

A Sports expenditure data

The expenditure data were purchased as Scientific Use Files (RDC of the Federal Statistical Office and Statistical Offices of the Länder, 2018b). Throughout our analysis, we focus on German municipalities (urban municipalities / cities and district municipalities). The reason for this is that expenditures by German states are usually channeled through municipality budgets. Municipalities represent the lowest level of public administration in Germany. Several preparation steps were necessary to obtain the sports expenditure measure we use throughout the analysis and we will outline these steps briefly in the following. A detailed description of all data preparation steps is available in Pawlowski et al. (2019), Appendix A.

The manual preparation of the data set encompassed among other things updating of the municipality identifiers and taking account for mergers of municipalities for all years. In the process of data preparation, the expenditure data were also corrected for World Cup 2006 expenditures (German Federal Government, 2006).

Since many communities outsource the operation of swimming pools as well as sports facilities into owner-operated community enterprises, we account for these as well in our analysis. We added expenditures on owner-operated community enterprises corrected for public subsidies (which are already included in SUF data) via an on-site workplace at a German regional data center (RDC of the Federal Statistical Office and Statistical Offices of the Länder, 2018a). Finally, numbers were converted to 2004 EUR using the Consumer Price Index from the Federal Statistical Office (2017) in Germany. The analysis is carried out for net values (expenditure grouping number 999 minus income grouping number 999) which is a proxy for local subsidies coming from tax income.

B Construction of SOEP sample and descriptive statistics

Steps in sample construction (Numbers are for main sample):

- Valid interviews in 2001 and 2008
- Individuals aged 17-65 years in 2001 & official municipality key not missing
- Exclude individuals from Berlin, Hamburg, Bremen (-654 observations)
- Exclude individuals with missings in individual or regional characteristics
- Restrict expenditure variable to positive values and up to the 95th percentile
- Exclude individuals who are still pursuing university or vocational education in 2001[47] (-62 observations)

Final sample sizes:

- A1 sample (aged 17-30 years): N = 2,015
- A2 sample (aged 31-45 years): N = 3,719
- A3 sample (aged 46-65 years): N = 4,007

[47] Note that we do not impose this restriction for the A1-sample.

Table B.1: Descriptive Statistics

	A1 17-30 years				A2 31-45 years				A3 46-65 years			
	T low	T middle	T high	Total	T low	T middle	T high	Total	T low	T middle	T high	Total
ln(Population per km^2/100)	1.208	1.436	1.739	1.448	1.027	1.355	1.806	1.398	1.158	1.310	1.843	1.436
	(1.316)	(1.162)	(1.127)	(1.227)	(1.275)	(1.133)	(1.074)	(1.206)	(1.331)	(1.109)	(1.026)	(1.197)
ln(Average rain/100)	1.994	2.009	2.101	2.032	2.032	2.036	2.120	2.063	1.998	2.019	2.110	2.042
	(0.212)	(0.217)	(0.236)	(0.225)	(0.214)	(0.213)	(0.220)	(0.219)	(0.220)	(0.210)	(0.223)	(0.223)
ln(Prop. recreational area pc in sqm/100)	-0.734	-0.821	-1.045	-0.859	-0.693	-0.856	-1.124	-0.892	-0.674	-0.796	-1.087	-0.852
	(0.814)	(0.711)	(0.735)	(0.766)	(0.766)	(0.727)	(0.670)	(0.743)	(0.835)	(0.750)	(0.686)	(0.778)
Unemployment mean	9.973	8.535	7.517	8.732	9.201	8.533	7.014	8.244	10.04	8.713	7.521	8.752
	(4.547)	(3.736)	(3.866)	(4.196)	(4.347)	(3.903)	(3.257)	(3.964)	(4.398)	(3.994)	(3.817)	(4.198)
Commuters mean	-39.13	-34.38	-19.17	-31.42	-53.79	-41.33	-16.53	-37.12	-39.73	-44.77	-17.01	-34.00
	(87.00)	(89.46)	(68.76)	(83.12)	(98.58)	(93.52)	(61.27)	(87.33)	(82.66)	(91.04)	(57.44)	(79.47)
Distance-weighted sports facility exp. 2002	10.33	28.06	43.26	26.45	9.505	26.40	43.25	26.49	10.44	27.00	43.70	27.10
	(18.50)	(13.42)	(17.48)	(21.30)	(17.05)	(12.92)	(17.98)	(21.18)	(15.50)	(11.09)	(18.35)	(20.30)
Male	0.436	0.428	0.461	0.441	0.471	0.481	0.477	0.476	0.480	0.490	0.465	0.479
Age in years	24.44	24.48	24.05	24.33	38.21	38.19	37.78	38.06	55.16	55.00	55.25	55.13
	(3.988)	(4.034)	(4.021)	(4.015)	(4.248)	(4.247)	(4.061)	(4.189)	(5.805)	(6.094)	(5.703)	(5.872)
Married	0.217	0.226	0.251	0.230	0.728	0.745	0.687	0.720	0.801	0.806	0.802	0.803
Divorced	0.015	0.013	0.006	0.011	0.092	0.084	0.085	0.087	0.080	0.088	0.088	0.085
German nationality	0.950	0.908	0.893	0.918	0.964	0.949	0.907	0.940	0.967	0.971	0.934	0.958
Number of children	0.546	0.582	0.522	0.551	1.252	1.285	1.139	1.225	0.219	0.246	0.237	0.234
	(0.821)	(0.828)	(0.788)	(0.813)	(1.058)	(1.060)	(1.050)	(1.057)	(0.552)	(0.618)	(0.586)	(0.587)
Lower secondary school/no degree (R)	0.212	0.192	0.216	0.207	0.283	0.256	0.261	0.267	0.482	0.515	0.486	0.495
Other graduation diploma	0.020	0.041	0.017	0.026	0.034	0.055	0.056	0.049	0.042	0.045	0.075	0.054
Intermediate secondary school	0.421	0.374	0.297	0.367	0.477	0.453	0.400	0.443	0.274	0.274	0.253	0.267
Upper secondary school	0.272	0.313	0.354	0.311	0.205	0.237	0.283	0.242	0.201	0.166	0.186	0.184
Not yet graduated	0.075	0.080	0.115	0.089	0	0	0	0	0	0	0	0
In education	0.264	0.244	0.291	0.265	0.017	0.015	0.021	0.018	0.006	0.002	0.002	0.003
No vocational degree	0.182	0.200	0.222	0.200	0.091	0.091	0.113	0.098	0.150	0.124	0.148	0.140
Degree below university	0.566	0.533	0.499	0.534	0.812	0.797	0.746	0.785	0.714	0.769	0.724	0.736
University degree	0.065	0.108	0.063	0.079	0.195	0.204	0.220	0.206	0.248	0.213	0.216	0.225
Northern states	0.150	0.154	0.023	0.112	0.178	0.135	0.035	0.116	0.184	0.143	0.031	0.120
Western/central states	0.060	0.126	0.179	0.119	0.107	0.141	0.237	0.162	0.071	0.140	0.190	0.134
Southern states	0.180	0.195	0.427	0.260	0.198	0.171	0.378	0.249	0.135	0.172	0.364	0.223
Eastern states	0.419	0.323	0.219	0.325	0.329	0.296	0.160	0.261	0.431	0.334	0.203	0.323
North Rhine-Westphalia (R)	0.192	0.203	0.153	0.184	0.189	0.257	0.190	0.212	0.179	0.210	0.211	0.200
Sports at least monthly 2001	0.446	0.492	0.542	0.491	0.363	0.375	0.406	0.381	0.315	0.290	0.377	0.327
Sports participation 2008	2.800	2.828	2.862	2.829	2.687	2.728	2.826	2.747	2.509	2.519	2.588	2.539
Sports at least monthly 2008	0.526	0.538	0.568	0.543	0.487	0.499	0.548	0.511	0.434	0.440	0.459	0.445

"Northern states": Schleswig-Holstein, Lower Saxony; *"West/central states"*: Hesse, Saarland, Rhineland-Palatinate; *"Eastern states"*: Mecklenburg-Vorpommern, Brandenburg, Saxony-Anhalt, Saxony, Thuringia; *"Southern states"*: Baden-Wuerttemberg, Bavaria; *Note:* Standard deviation in parenthesis. The mean of the variable "not yet graduated" is zero in the A2 and A3 sample, since all individuals graduated from school by the age of 31. *Sources:* RDC of the Federal Statistical Office and Statistical Offices of the Länder (2018a; 2018b), SOEP as well as further data sources as discussed in Chapter 2, own calculations.

C GPS-estimation, common support & balancing tests

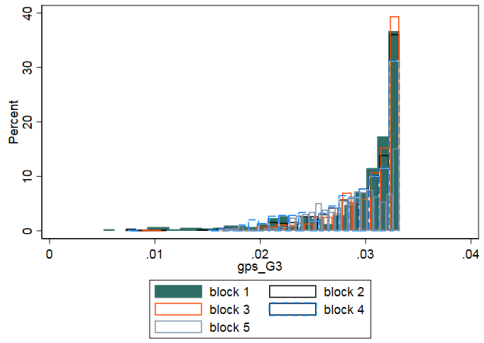
Table C.1: Estimation of GPS

VARIABLES	Treatment
ln(population per km^2/100)	0.0830***
	(0.0265)
ln(average rain/100)	0.345***
	(0.118)
ln(proportion recreational area p.c. in sqm/100)	0.00381
	(0.0315)
Unemployment mean	-0.0400***
	(0.0104)
Commuters mean	0.000768***
	(0.000189)
Age in years	-0.00256
	(0.00235)
Number of children	-0.0132
	(0.0111)
Male	-0.00312
	(0.0126)
Married	0.0198
	(0.0284)
Divorced	0.0112
	(0.0367)
German nationality	-0.0899**
	(0.0434)
Other graduation diploma	0.101**
	(0.0427)
Intermediate secondary school	0.0575**
	(0.0238)
Upper secondary school	0.126***
	(0.0330)
No vocational degree	0.0705
	(0.0477)
Degree below university	0.0579
	(0.0400)
University degree	-0.0238
	(0.0361)
Northern states	-0.184**
	(0.0745)
Western/central states	0.239***
	(0.0689)
Southern states	0.172***
	(0.0628)
Eastern states	0.262**
	(0.123)
Observations	3,719

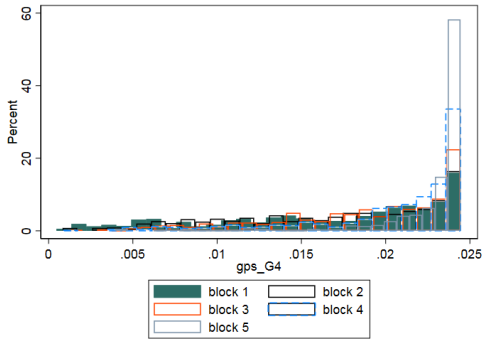
Coefficients of the GPS Model, constant included.

Sources: RDC of the Federal Statistical Office and Statistical Offices of the Länder (2018a; 2018b), SOEP as well as further data sources as discussed in Chapter 2, own calculations.

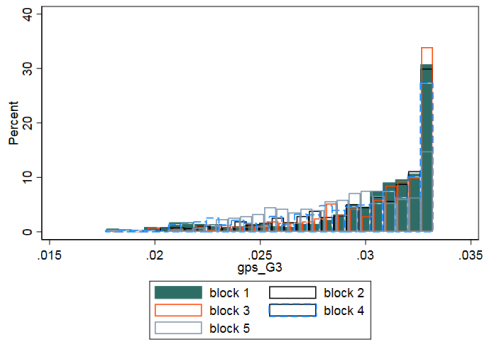
Figure C.1: Common support - A2 sample (main sample)



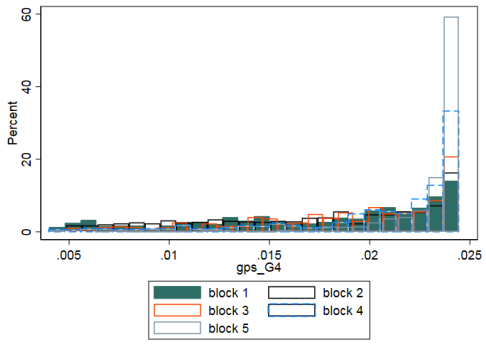
(a) GPS at p50 - before CS condition



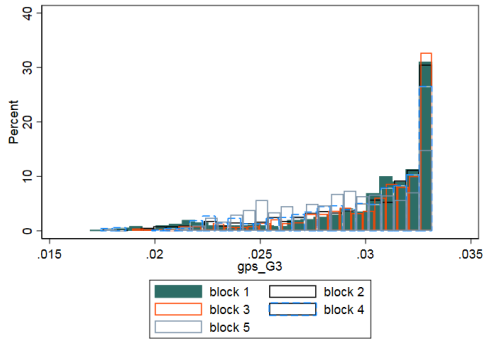
(b) GPS at p70 - before CS condition



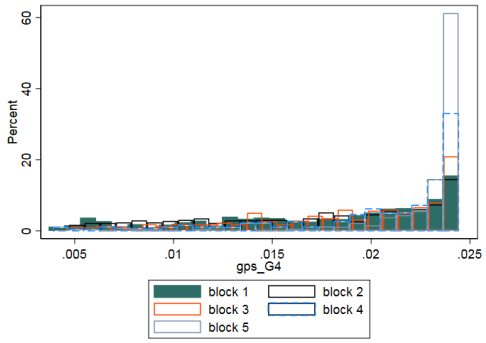
(c) GPS at p50 - after step 1



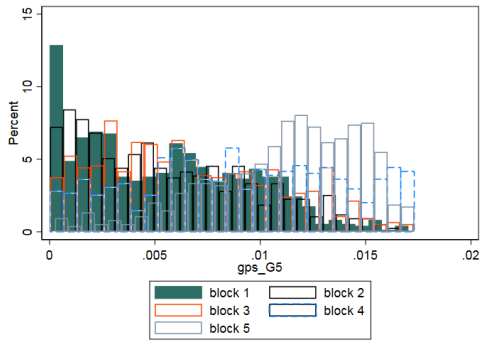
(d) GPS at p70 - after step 1



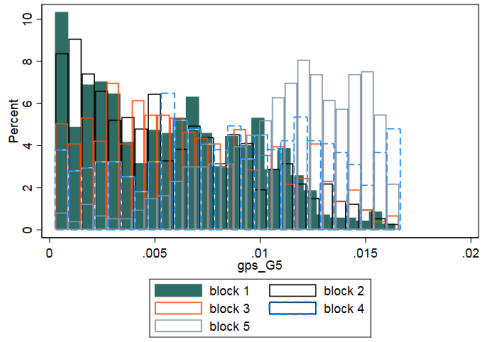
(e) GPS at p50 - after step 3



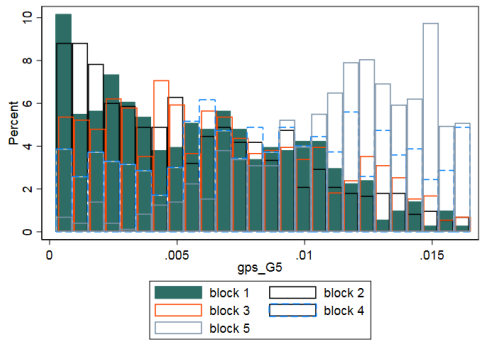
(f) GPS at p70 - after step 3



(g) GPS at p90 - before CS condition



(h) GPS at p90 - after round 1



(i) GPS at p90 - after round 3

Sources: RDC of the Federal Statistical Office and Statistical Offices of the Länder (2018a; 2018b), SOEP as well as further data sources as discussed in Chapter 2, own calculations.

Table C.2: Balancing before GPS adjustment

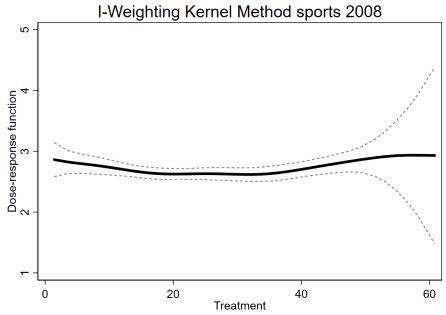
	Block 1 vs 2			Block 1 vs 3			Block 1 vs 4			Block 1 vs 5			Block 2 vs 3		
	Mean 1	Mean 2	% bias	Mean 1	Mean 3	% bias	Mean 1	Mean 4	% bias	Mean 1	Mean 5	% bias	Mean 2	Mean 3	% bias
ln(population per km2/100)	1.15	1.07	6.54	1.15	1.37	-18.04	1.15	1.71	-43.80	1.15	1.81	-56.21	1.07	1.37	-26.17
ln(average rain/100)	2.02	2.02	1.94	2.02	2.05	-11.02	2.02	2.09	-29.50	2.02	2.14	-53.97	2.02	2.05	-13.28
ln(proportion recreational area pc in sqm/100)	-0.68	-0.73	7.12	-0.68	-0.83	20.67	-0.68	-1.03	51.62	-0.68	-1.17	71.48	-0.73	-0.83	11.81
Unemployment mean	9.26	9.07	4.41	9.26	8.31	22.92	9.26	7.79	37.27	9.26	6.40	78.06	9.07	8.31	18.77
Commuters mean	-45.12	-50.32	5.60	-45.12	-42.58	-2.76	-45.12	-20.12	-31.18	-45.12	-16.67	-36.85	-50.32	-42.58	-8.18
Age	38.03	38.36	-7.81	38.03	38.08	-1.27	38.03	37.85	4.27	38.03	37.73	6.92	38.36	38.08	6.57
Number of children	1.27	1.32	-5.05	1.27	1.32	-4.81	1.27	1.27	-0.17	1.27	1.23	3.27	1.32	1.32	0.09
Male	0.47	0.46	0.01	0.47	0.49	-0.02	0.47	0.48	-0.01	0.47	0.47	0.00	0.46	0.49	-0.03
Married	0.71	0.74	-0.03	0.71	0.76	-0.05	0.71	0.72	-0.01	0.71	0.73	-0.02	0.74	0.76	-0.02
Divorced	0.09	0.09	0.00	0.09	0.08	0.01	0.09	0.08	0.01	0.09	0.08	0.01	0.09	0.08	0.02
German	0.95	0.96	0.00	0.95	0.94	0.02	0.95	0.90	0.05	0.95	0.87	0.08	0.96	0.94	0.02
Other graduation diploma	0.04	0.05	-0.01	0.04	0.05	-0.01	0.04	0.07	-0.03	0.04	0.09	-0.05	0.05	0.05	0.00
Intermediate secondary school	0.45	0.49	-0.04	0.45	0.43	0.02	0.45	0.41	0.04	0.45	0.34	0.11	0.49	0.43	0.06
Upper secondary school	0.21	0.19	0.01	0.21	0.24	-0.03	0.21	0.24	-0.04	0.21	0.28	-0.08	0.19	0.24	-0.05
No vocational degree	0.11	0.10	0.02	0.11	0.10	0.01	0.11	0.12	-0.01	0.11	0.16	-0.04	0.10	0.10	-0.01
Degree below university	0.79	0.80	-0.01	0.79	0.78	0.00	0.79	0.75	0.03	0.79	0.71	0.07	0.80	0.78	0.02
University	0.18	0.22	-0.04	0.18	0.17	0.01	0.18	0.20	-0.02	0.18	0.22	-0.04	0.22	0.17	0.05
Northern states	0.21	0.16	0.05	0.21	0.15	0.06	0.21	0.08	0.13	0.21	0.02	0.19	0.16	0.15	0.01
Western states	0.11	0.09	0.02	0.11	0.16	-0.05	0.11	0.16	-0.05	0.11	0.24	-0.12	0.09	0.16	-0.07
Southern states	0.18	0.20	-0.03	0.18	0.18	0.00	0.18	0.33	-0.15	0.18	0.43	-0.25	0.20	0.18	0.02
Eastern states	0.30	0.36	-0.05	0.30	0.25	0.05	0.30	0.22	0.08	0.30	0.09	0.21	0.36	0.25	0.11

	Block 2 vs 4			Block 2 vs 5			Block 3 vs 4			Block 3 vs 5			Block 4 vs 5		
	Mean 2	Mean 4	% bias	Mean 2	Mean 5	% bias	Mean 3	Mean 4	% bias	Mean 3	Mean 5	% bias	Mean 4	Mean 5	% bias
ln(population per km2/100)	1.07	1.71	-52.89	1.07	1.81	-67.29	1.37	1.71	-28.85	1.37	1.81	-41.39	1.71	1.81	-8.70
ln(average rain/100)	2.02	2.09	-31.78	2.02	2.14	-56.98	2.05	2.09	-20.52	2.05	2.14	-45.51	2.09	2.14	-20.26
ln(proportion recreational area pc in sqm/100)	-0.73	-1.03	40.01	-0.73	-1.17	58.26	-0.83	-1.03	31.12	-0.83	-1.17	51.45	-1.03	-1.17	20.93
Unenployment mean	9.07	7.79	33.22	9.07	6.40	74.75	8.31	7.79	13.76	8.31	6.40	54.91	7.79	6.40	43.05
Commuters mean	-50.32	-20.12	-36.18	-50.32	-16.67	-41.74	-42.58	-20.12	-27.41	-42.58	-16.67	-32.78	-20.12	-16.67	-5.27
Age	38.36	37.85	12.53	38.36	37.73	14.97	38.08	37.85	5.63	38.08	37.73	8.25	37.85	37.73	2.90
Number of children	1.32	1.27	4.91	1.32	1.23	8.46	1.32	1.27	4.67	1.32	1.23	8.12	1.27	1.23	3.47
Male	0.46	0.48	-0.02	0.46	0.47	-0.01	0.49	0.48	0.01	0.49	0.47	0.02	0.48	0.47	0.01
Married	0.74	0.72	0.02	0.74	0.73	0.02	0.76	0.72	0.04	0.76	0.73	0.04	0.72	0.73	0.00
Divorced	0.09	0.08	0.01	0.09	0.08	0.01	0.08	0.08	0.00	0.08	0.08	0.00	0.08	0.08	0.00
German	0.96	0.90	0.05	0.96	0.87	0.09	0.94	0.90	0.03	0.94	0.87	0.06	0.90	0.87	0.03
Other graduation diploma	0.05	0.07	-0.02	0.05	0.09	-0.04	0.05	0.07	-0.02	0.05	0.09	-0.04	0.07	0.09	-0.03
Intermediate secondary school	0.49	0.41	0.08	0.49	0.34	0.15	0.43	0.41	0.02	0.43	0.34	0.09	0.41	0.34	0.07
Upper secondary school	0.19	0.24	-0.05	0.19	0.28	-0.09	0.24	0.24	0.00	0.24	0.28	-0.04	0.24	0.28	-0.04
No vocational degree	0.10	0.12	-0.03	0.10	0.16	-0.06	0.10	0.12	-0.02	0.10	0.16	-0.05	0.12	0.16	-0.03
Degree below university	0.80	0.75	0.04	0.80	0.71	0.09	0.78	0.75	0.03	0.78	0.71	0.07	0.75	0.71	0.04
University	0.22	0.20	0.02	0.22	0.22	0.00	0.17	0.20	-0.03	0.17	0.22	-0.05	0.20	0.22	-0.01
Northern states	0.16	0.08	0.08	0.16	0.02	0.14	0.15	0.08	0.07	0.15	0.02	0.14	0.08	0.02	0.07
Western states	0.09	0.16	-0.07	0.09	0.24	-0.15	0.16	0.16	0.00	0.16	0.24	-0.08	0.16	0.24	-0.07
Southern states	0.20	0.33	-0.12	0.20	0.43	-0.23	0.18	0.33	-0.14	0.18	0.43	-0.25	0.33	0.43	-0.10
Eastern states	0.36	0.22	0.14	0.36	0.09	0.27	0.25	0.22	0.03	0.25	0.09	0.16	0.22	0.09	0.13

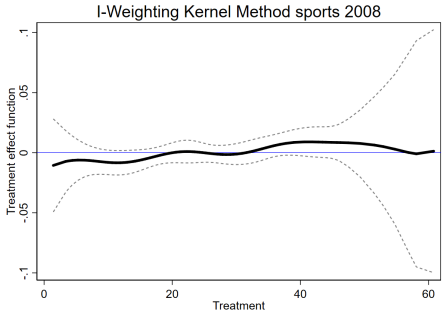
Sources: RDC of the Federal Statistical Office and Statistical Offices of the Länder (2018a; 2018b), SOEP as well as further data sources as discussed in Chapter 2, own calculations.

D Robustness checks

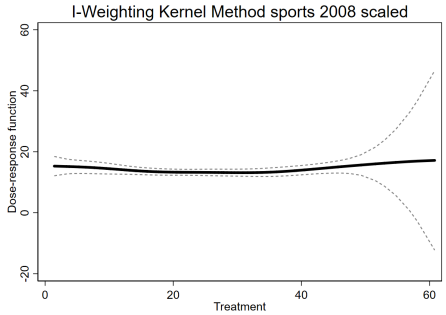
Figure D.1: Results different specifications outcome variable



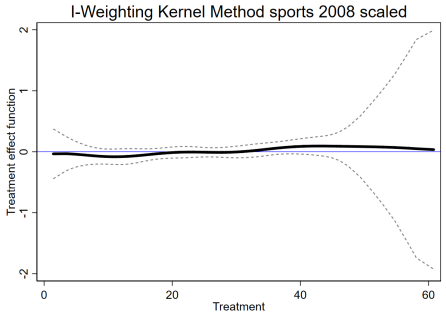
(a) DRF sports 'raw' variable - 2008



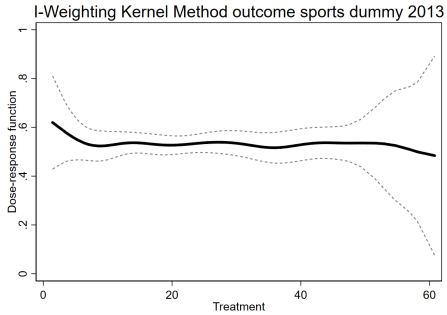
(b) TEF sports 'raw' variable - 2008



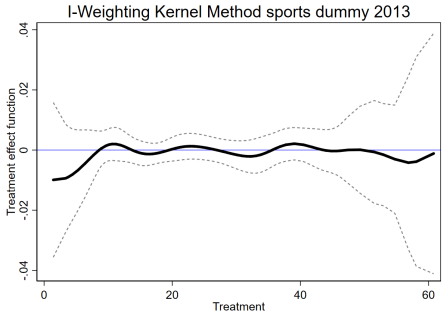
(c) DRF sports variable scaled - 2008



(d) TEF sports variable scaled - 2008



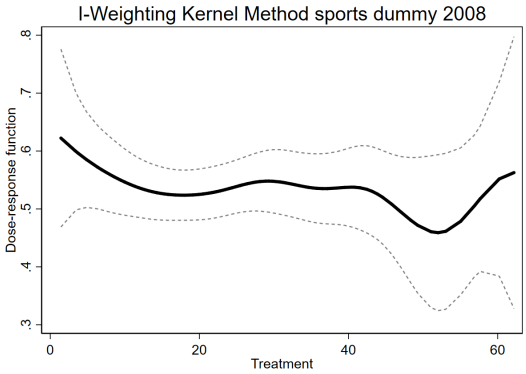
(e) DRF sports dummy - 2013



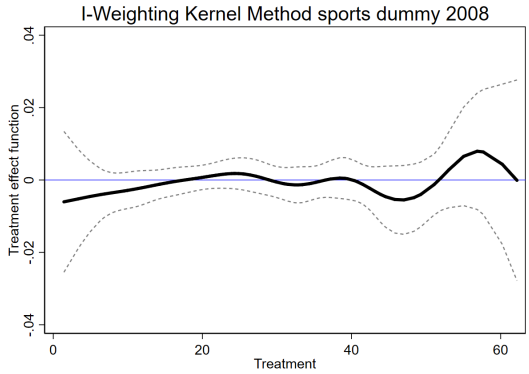
(f) TEF sports dummy - 2013

Note: "DRF": dose-response function; *"TEF"*: treatment effect function; *"I-weighting Kernel method"* denotes the Inverse-weighting Kernel method as proposed by Flores et al. (2012). *Sources:* RDC of the Federal Statistical Office and Statistical Offices of the Länder (2018a; 2018b), SOEP as well as further data sources as discussed in Chapter 2, own calculations.

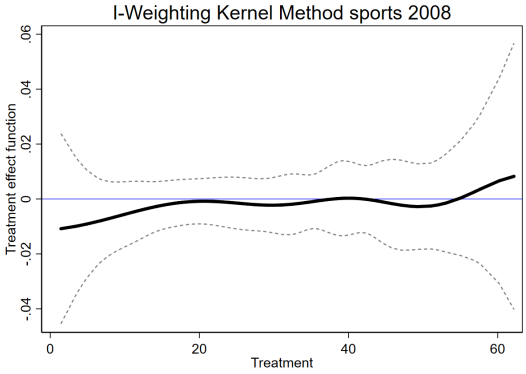
Figure D.2: Results A1 sample (17-30 years of age)



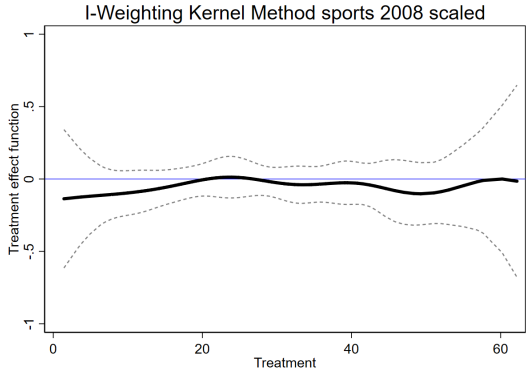
(a) DRF sports dummy - 2008



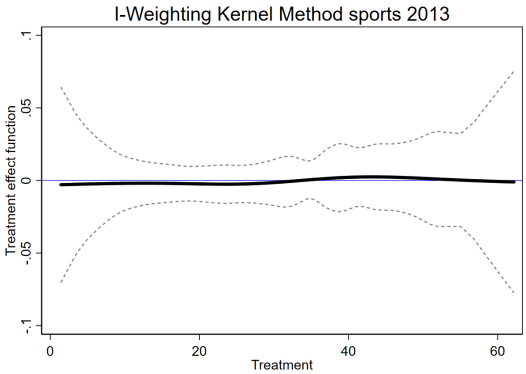
(b) TEF sports dummy - 2008



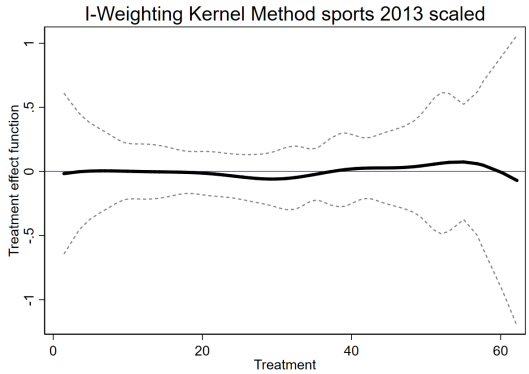
(c) TEF sports 'raw' variable - 2008



(d) TEF sports variable scaled - 2008



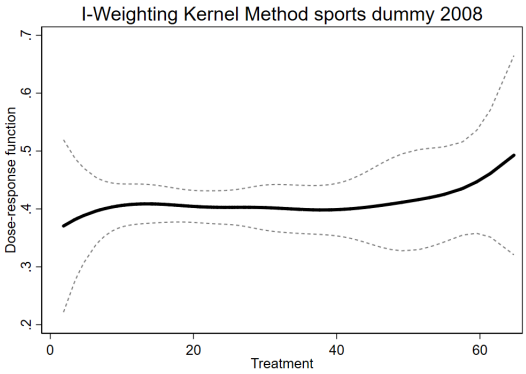
(e) TEF sports 'raw' variable - 2013



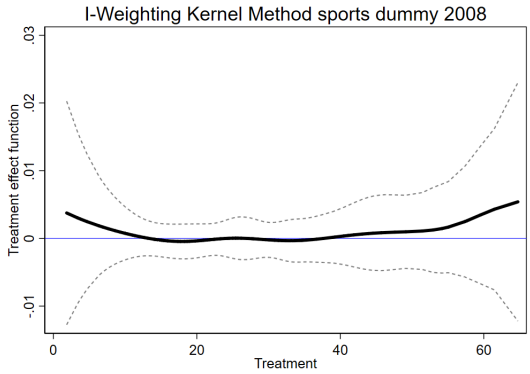
(f) TEF sports variable scaled - 2013

Sources: RDC of the Federal Statistical Office and Statistical Offices of the Länder (2018a; 2018b), SOEP as well as further data sources as discussed in Chapter 2, own calculations.

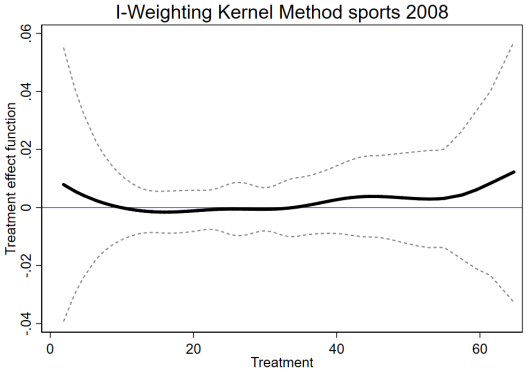
Figure D.3: Results A3 sample (46-65 years of age)



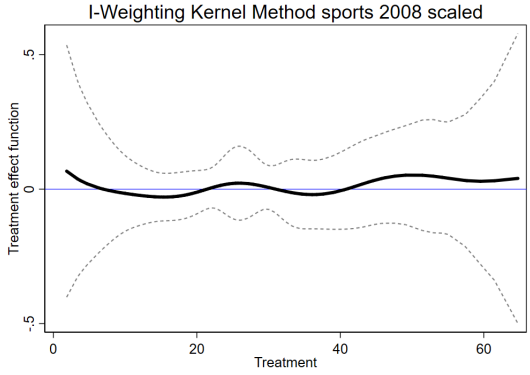
(a) DRF sports dummy - 2008



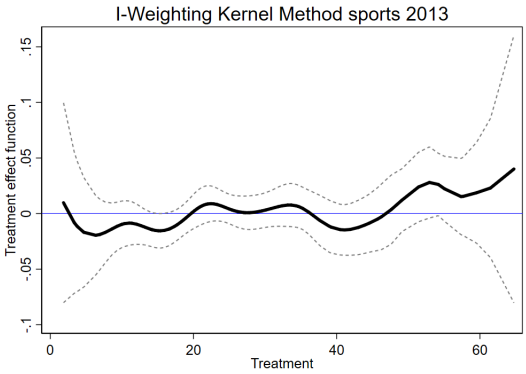
(b) TEF sports dummy - 2008



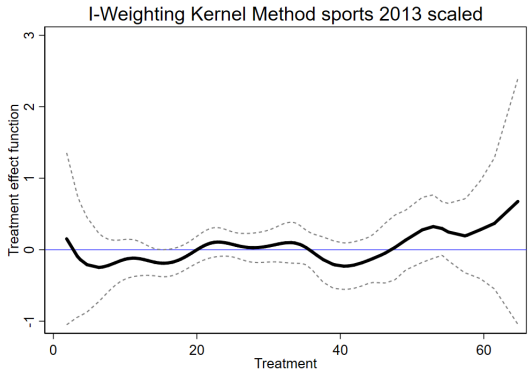
(c) TEF sports 'raw' variable - 2008



(d) TEF sports variable scaled - 2008



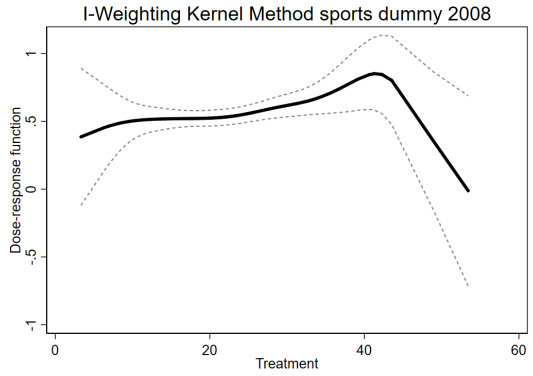
(e) TEF sports 'raw' variable - 2013



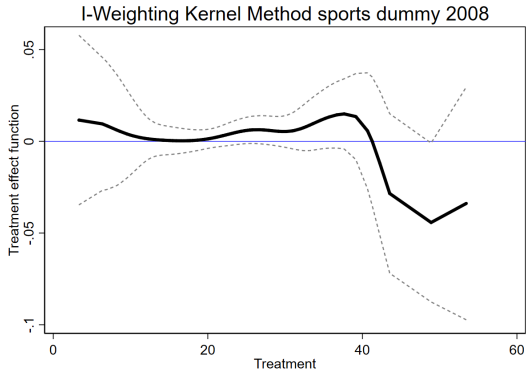
(f) TEF sports variable scaled - 2013

Sources: RDC of the Federal Statistical Office and Statistical Offices of the Länder (2018a; 2018b), SOEP as well as further data sources as discussed in Chapter 2, own calculations.

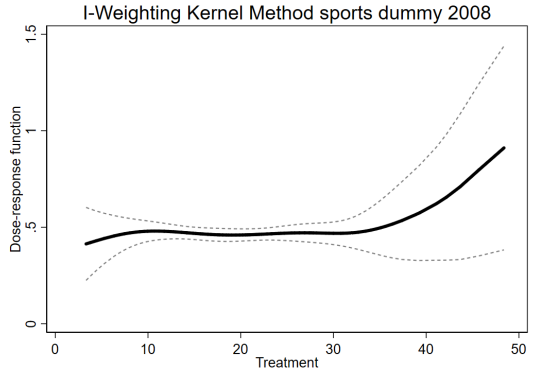
Figure D.4: Results with lagged outcome and treatment



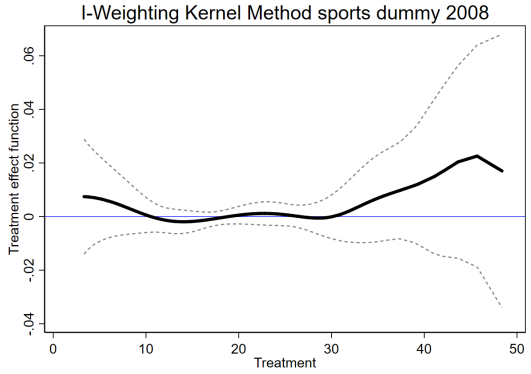
(a) DRF sports dummy 2008 - A1 sample



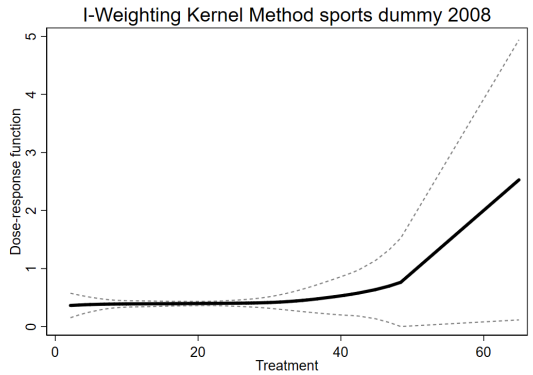
(b) TEF sports dummy 2008 - A1 sample



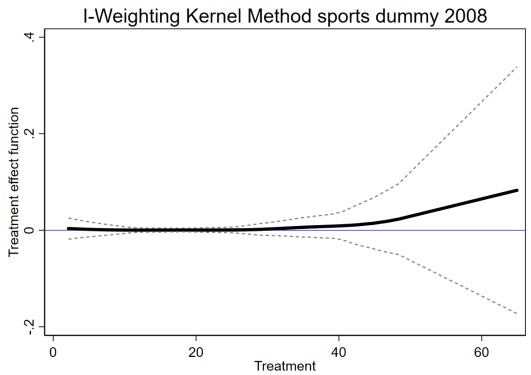
(c) DRF sports dummy 2008 - A2 sample



(d) TEF sports dummy 2008 - A2 sample



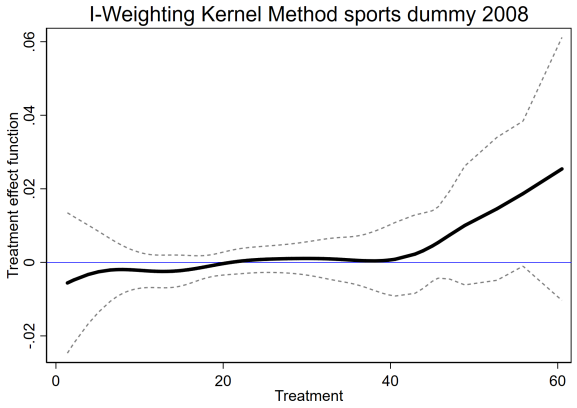
(e) DRF sports dummy 2008 - A3 sample



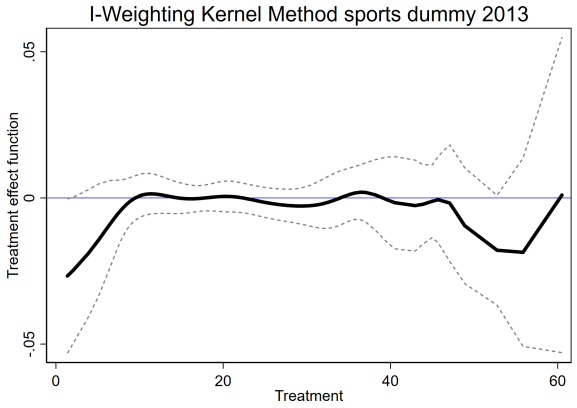
(f) TEF sports dummy 2008 - A3 sample

Sources: RDC of the Federal Statistical Office and Statistical Offices of the Länder (2018a; 2018b), SOEP as well as further data sources as discussed in Chapter 2, own calculations.

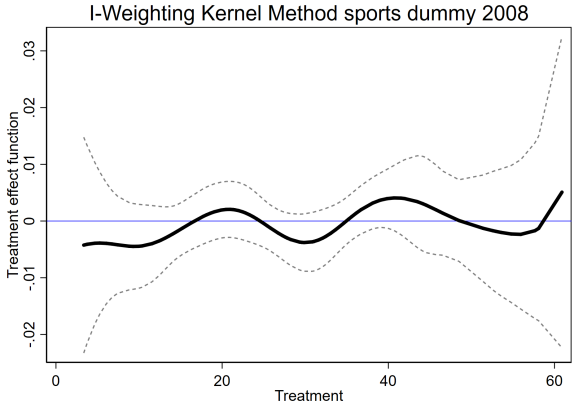
Figure D.5: Results small / large municipalities



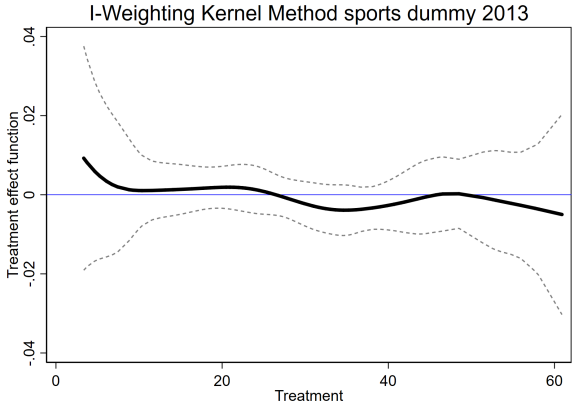
(a) Sports dummy 2008 - Small municipalities



(b) Sports dummy 2013 - Small municipalities



(c) Sports dummy 2008 - Large municipalities



(d) Sports dummy 2013 - Large municipalities

Sources: RDC of the Federal Statistical Office and Statistical Offices of the Länder (2018a; 2018b), SOEP as well as further data sources as discussed in Chapter 2, own calculations.

E Additional estimation results

1. **GPS Estimation Results**
   - A1 sample (17-30 years)
   - A3 sample (46-65 years)
   - A1 with lagged outcome and treatment
   - A2 with lagged outcome and treatment
   - A3 with lagged outcome and treatment
   - Small municipalities
   - Large municipalities

2. **Common Support Histograms**
   - A1 sample (17-30 years)
   - A3 sample (46-65 years)
   - A1 with lagged outcome and treatment
   - A2 with lagged outcome and treatment
   - A3 with lagged outcome and treatment
   - Small municipalities
   - Large municipalities

3. **Balancing**
   - A1 sample (17-30 years)
   - A3 sample (46-65 years)
   - A1 with lagged outcome and treatment
   - A2 with lagged outcome and treatment
   - A3 with lagged outcome and treatment
   - Small municipalities
   - Large municipalities

4. **Dose-response functions**
   - A1 sample (17-30 years)
   - A3 sample (46-65 years)

Table E.1: Estimation of GPS - A1 sample (17-30 years)

VARIABLES	Treatment
ln(Population per km^2/100)	0.0744**
	(0.0300)
ln(Average rain/100)	0.512***
	(0.141)
ln(Proportion recreational area p.c. in sqm/100)	0.0719**
	(0.0367)
Unemployment mean	-0.0276**
	(0.0119)
Commuters mean	0.000479*
	(0.000261)
Age in years	-0.00337
	(0.00416)
Number of children	-0.00846
	(0.0189)
Male	0.0352*
	(0.0200)
Married	0.0397
	(0.0411)
Divorced	-0.0106
	(0.108)
German nationality	-0.0764
	(0.0480)
Other graduation diploma	0.0540
	(0.0647)
Intermediate secondary school	0.0596
	(0.0369)
Upper secondary school	0.0951**
	(0.0425)
Not yet graduated	0.0783
	(0.0591)
In education	-0.0366
	(0.0523)
No vocational degree	-0.0327
	(0.0604)
degree below university	-0.0467
	(0.0514)
University degree	-0.0139
	(0.0525)
Northern states	-0.207**
	(0.0883)
Western/central states	0.325***
	(0.0827)
Southern states	0.263***
	(0.0773)
Eastern states	0.221
	(0.141)
Constant	2.370***
	(0.336)
Observations	2,015

Coefficients of the GPS Model

Sources: RDC of the Federal Statistical Office and Statistical Offices of the Länder (2018a; 2018b), SOEP as well as further data sources as discussed in Chapter 2, own calculations.

Table E.2: Estimation of GPS - A3 sample (46-65 years)

VARIABLES	Treatment
ln(Population per km^2/100)	0.0842**
	(0.0364)
ln(Average rain/100)	0.343**
	(0.159)
ln(Proportion recreational area p.c. in sqm/100)	0.0362
	(0.0409)
Unemployment mean	-0.0343***
	(0.0124)
Commuters mean	0.000753***
	(0.000230)
Age in years	0.00116
	(0.00162)
Number of children	-0.00996
	(0.0165)
Male	-0.0164
	(0.0112)
Married	0.0302
	(0.0271)
Divorced	0.00688
	(0.0379)
German nationality	-0.0804
	(0.0504)
Other graduation diploma	-0.0208
	(0.0516)
Intermediate secondary school	-0.00575
	(0.0255)
Upper secondary school	0.00252
	(0.0450)
No vocational degree	-0.0177
	(0.0512)
Degree below university	0.0277
	(0.0407)
University degree	-0.00929
	(0.0325)
Northern states	-0.264***
	(0.0947)
Western/central states	0.237***
	(0.0814)
Southern states	0.171**
	(0.0746)
Eastern states	0.172
	(0.150)
Constant	2.666***
	(0.350)
Observations	4,007

Coefficients of the GPS Model

Sources: RDC of the Federal Statistical Office and Statistical Offices of the Länder (2018a; 2018b), SOEP as well as further data sources as discussed in Chapter 2, own calculations.

Table E.3: Estimation of GPS - A1 sample with lagged outcome & treatment

VARIABLES	Treatment
ln(Population per km^2/100)	0.0329
	(0.0221)
ln(Average rain/100)	0.288***
	(0.0921)
ln(Proportion recreational area pc in sqm/100)	0.0524*
	(0.0274)
Unemployment mean	-0.0216**
	(0.00962)
Commuters mean	0.000359*
	(0.000195)
Dist. weighted sports facility expenditures 2002	0.0167***
	(0.00173)
Age in years	-0.00511
	(0.00382)
Number of children	-0.00248
	(0.0165)
Male	0.0118
	(0.0206)
Married	0.0676*
	(0.0361)
Divorced	0.145
	(0.143)
German nationality	-0.0560
	(0.0451)
Other graduation diploma	-0.00125
	(0.0558)
Intermediate secondary school	0.0278
	(0.0309)
Upper secondary school	0.0411
	(0.0359)
Not yet graduated	0.0693
	(0.0541)
In education	-0.0120
	(0.0437)
No vocational degree	0.00111
	(0.0554)
Degree below university	-0.00511
	(0.0435)
University degree	0.0229
	(0.0441)
Northern states	-0.100
	(0.0685)
Western/central states	0.194***
	(0.0631)
Southern states	0.212***
	(0.0540)
Eastern states	0.201*
	(0.108)
Sports at least monthly 2001	0.00640
	(0.0233)
Constant	2.330***
	(0.243)
Observations	2,015

Coefficients of the GPS Model

Sources: RDC of the Federal Statistical Office and Statistical Offices of the Länder (2018a; 2018b), SOEP as well as further data sources as discussed in Chapter 2, own calculations.

Table E.4: Estimation of GPS - A2 sample with lagged outcome & treatment

VARIABLES	Treatment
ln(Population per km^2/100)	0.0418**
	(0.0208)
ln(Average rain/100)	0.176**
	(0.0844)
ln(Proportion recreational area p.c. in sqm/100)	-0.0117
	(0.0256)
Unemployment mean	-0.0236***
	(0.00838)
Commuters mean	0.000464***
	(0.000173)
Dist. weighted sports facility expenditures 2002	0.0166***
	(0.00159)
Age in years	-0.00192
	(0.00229)
Number of children	-0.000794
	(0.0101)
Male	-0.00794
	(0.0111)
Married	0.0354
	(0.0245)
Divorced	0.0356
	(0.0302)
German nationality	-0.113***
	(0.0433)
Other graduation diploma	0.0248
	(0.0342)
Intermediate secondary school	0.0624***
	(0.0200)
Upper sec. school	0.114***
	(0.0311)
No vocational degree	-0.000452
	(0.0424)
Degree below university	0.0358
	(0.0356)
University degree	-0.0249
	(0.0295)
Northern states	-0.108
	(0.0669)
Western/central states	0.128**
	(0.0543)
Southern states	0.122**
	(0.0521)
Eastern states	0.174*
	(0.101)
Sports at least monthly 2001	0.0175
	(0.0168)
Constant	2.487***
	(0.236)
Observations	3,719

Coefficients of the GPS Model

Sources: RDC of the Federal Statistical Office and Statistical Offices of the Länder (2018a; 2018b), SOEP as well as further data sources as discussed in Chapter 2, own calculations.

Table E.5: Estimation of GPS - A3 sample with lagged outcome & treatment

VARIABLES	Treatment
ln(population per km^2/100)	0.0447*
	(0.0260)
ln(average rain/100)	0.277**
	(0.126)
ln(proportion recreational area pc in sqm/100)	0.0262
	(0.0321)
Unemployment mean	-0.0272***
	(0.00920)
Commuters mean	0.000226
	(0.000204)
Dist. weighted sports facility expenditures 2002	0.0208***
	(0.00190)
Age in years	0.000294
	(0.00153)
Number of children	0.0223
	(0.0200)
Male	-0.0117
	(0.00898)
Married	0.0152
	(0.0241)
Divorced	0.0122
	(0.0388)
German nationality	-0.0844*
	(0.0460)
Other graduation diploma	-0.0282
	(0.0374)
Intermediate secondary school	-0.0247
	(0.0213)
Upper secondary school	0.00344
	(0.0363)
No vocational degree	-0.0102
	(0.0576)
Degree below university	0.0137
	(0.0490)
University degree	0.000879
	(0.0345)
Northern states	-0.0791
	(0.0724)
Western/central states	0.150**
	(0.0586)
Southern states	0.141***
	(0.0532)
Eastern states	0.275**
	(0.113)
Sports at least monthly 2001	0.0469**
	(0.0228)
Constant	2.150***
	(0.341)
Observations	4,007

Coefficients of the GPS Model

Sources: RDC of the Federal Statistical Office and Statistical Offices of the Länder (2018a; 2018b), SOEP as well as further data sources as discussed in Chapter 2, own calculations.

Table E.6: Estimation of GPS - A2 sample - small municipalities

VARIABLES	Treatment
ln(population per km^2/100)	0.147***
	(0.0252)
ln(average rain/100)	0.339***
	(0.116)
ln(proportion recreational area p.c. in sqm/100)	-0.00497
	(0.0305)
Unemployment mean	-0.0340***
	(0.0101)
Commuters mean	0.000767***
	(0.000197)
Age in years	-0.00485
	(0.00363)
Nr of children	-0.00546
	(0.0155)
Male	-0.0245
	(0.0184)
Married	0.0312
	(0.0461)
Divorced	0.00513
	(0.0595)
German nationality	-0.268***
	(0.0748)
Other graduation diploma	0.0628
	(0.0605)
Intermediate sec. school	0.0530
	(0.0339)
Upper sec. school	0.143***
	(0.0474)
No vocational degree	0.131*
	(0.0740)
degree below univ.	0.102*
	(0.0593)
University degree	-0.0406
	(0.0419)
Northern states	-0.112
	(0.105)
Western/central states	0.178*
	(0.0970)
Southern states	0.148*
	(0.0858)
Eastern states	0.337***
	(0.113)
Constant	2.817***
	(0.325)
Observations	1,782

Coefficients of the GPS Model

Sources: RDC of the Federal Statistical Office and Statistical Offices of the Länder (2018a; 2018b), SOEP as well as further data sources as discussed in Chapter 2, own calculations.

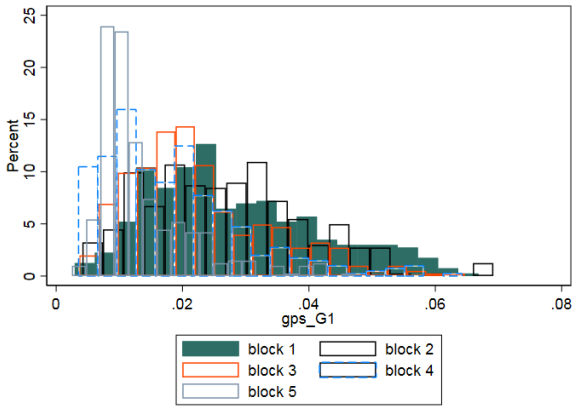
Table E.7: Estimation of GPS - A2 sample - large municipalities

VARIABLES	Treatment
ln(Population per km^2/100)	0.0246
	(0.0550)
ln(Average rain/100)	0.418**
	(0.210)
ln(Proportion recreational area p.c. in sqm/100)	0.0587
	(0.0754)
Unemployment mean	-0.0337*
	(0.0202)
Commuters mean	-0.000642
	(0.000708)
Age in years	-0.000270
	(0.00317)
Number of children	-0.0186
	(0.0162)
Male	0.0175
	(0.0168)
Married	-0.00114
	(0.0338)
Divorced	0.00358
	(0.0434)
German nationality	0.000836
	(0.0496)
Other graduation diploma	0.0865
	(0.0566)
Intermediate secondary school	0.0364
	(0.0322)
Upper secondary school	0.0937**
	(0.0448)
No vocational degree	0.0353
	(0.0617)
degree below university	0.0352
	(0.0541)
University degree	-0.0115
	(0.0597)
Northern states	-0.259***
	(0.0990)
Western/central states	0.393***
	(0.108)
Southern states	0.301***
	(0.0993)
Eastern states	0.115
	(0.223)
Constant	2.558***
	(0.524)
Observations	1,937

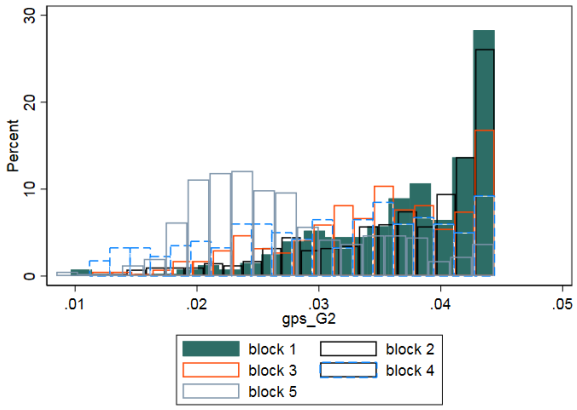
Coefficients of the GPS Model

Sources: RDC of the Federal Statistical Office and Statistical Offices of the Länder (2018a; 2018b), SOEP as well as further data sources as discussed in Chapter 2, own calculations.

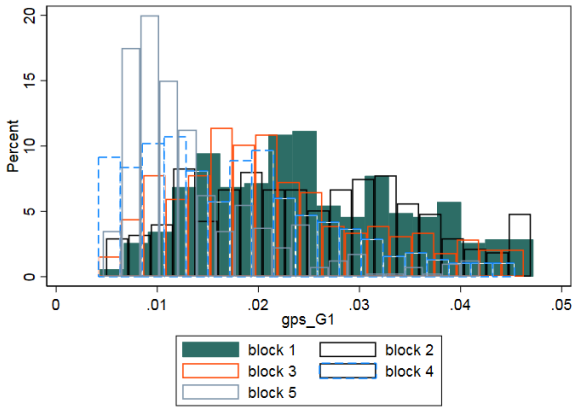
Figure E.1: Common support - A1 sample (17-30 years)



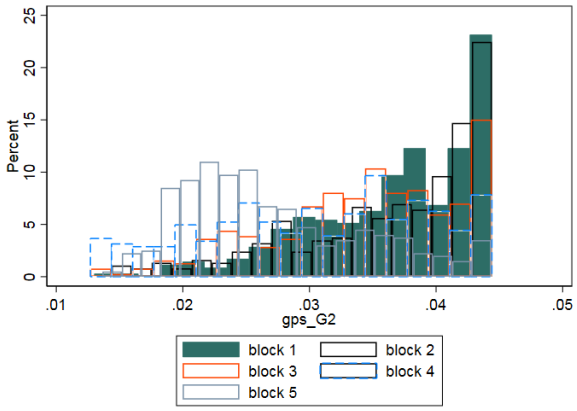
(a) GPS at p10 - before CS condition



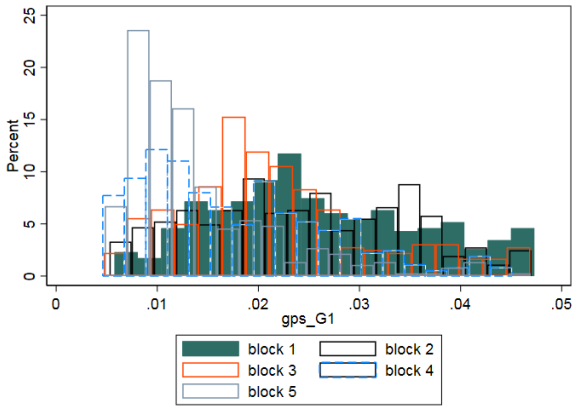
(b) GPS at p30 - before CS condition



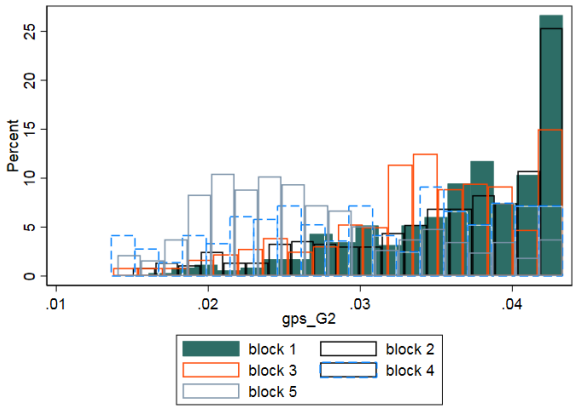
(c) GPS at p10 - after step 1



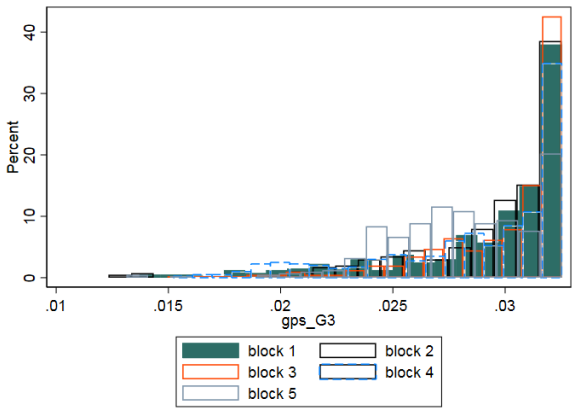
(d) GPS at p30 - after step 1



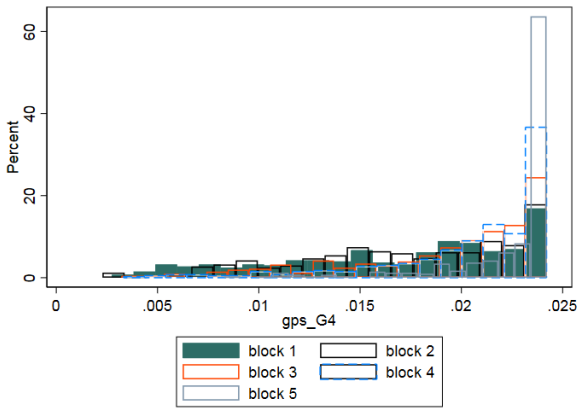
(e) GPS at p10 - after step 3



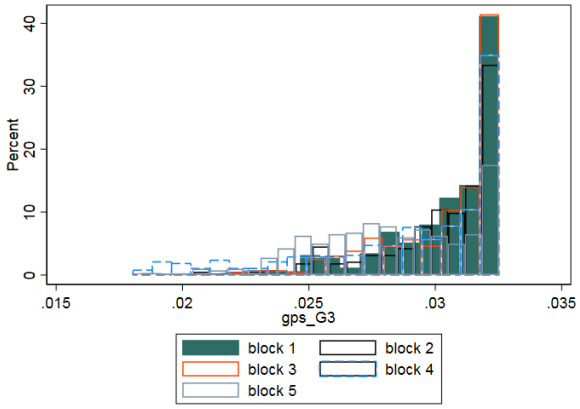
(f) GPS at p30 - after step 3



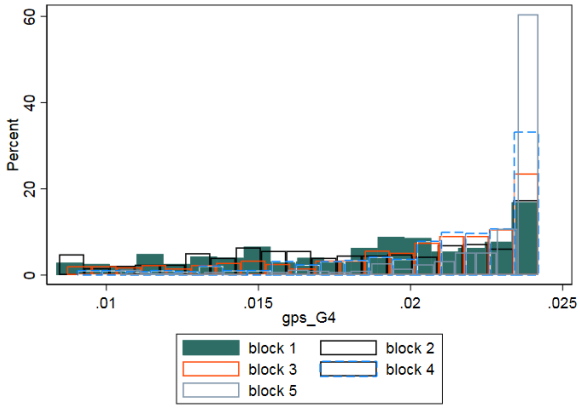
(g) GPS at p50 - before CS condition



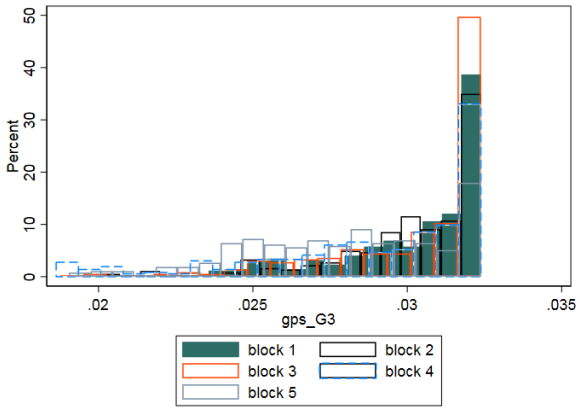
(h) GPS at p70 - before CS condition



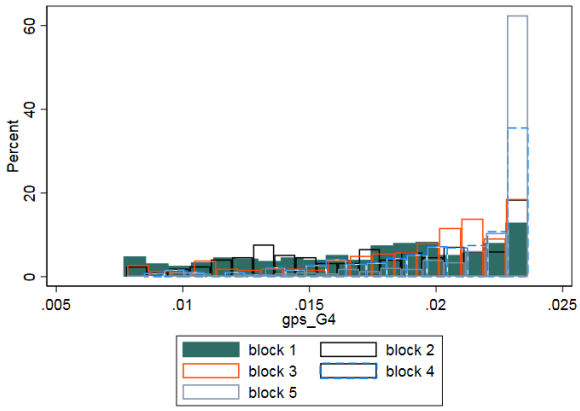
(i) GPS at p50 - after step 1



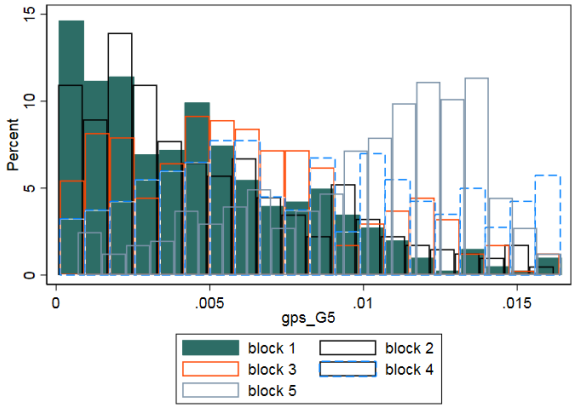
(j) GPS at p70 - after step 1



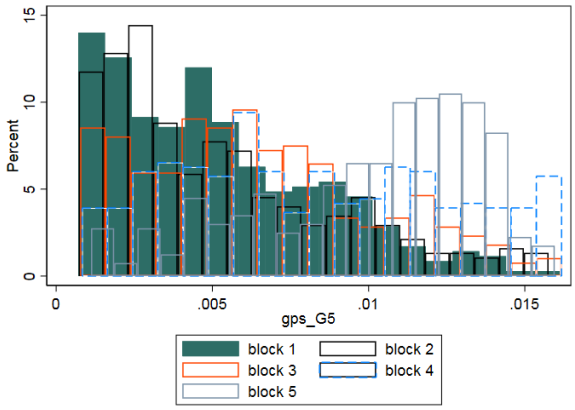
(k) GPS at p50 - after step 3



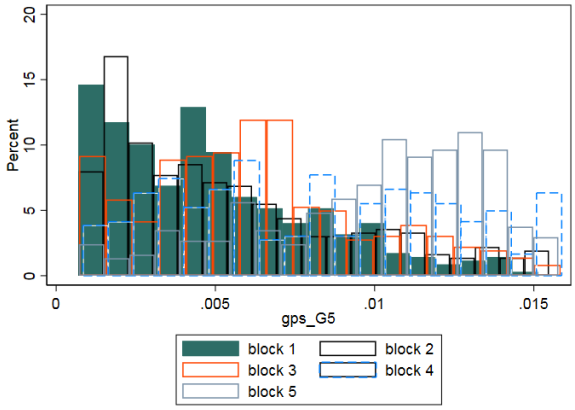
(l) GPS at p70 - after step 3



(m) GPS at p90 - before CS condition



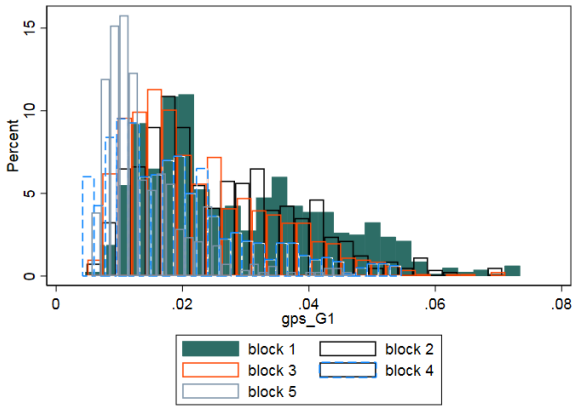
(n) GPS at p90 - after step 1



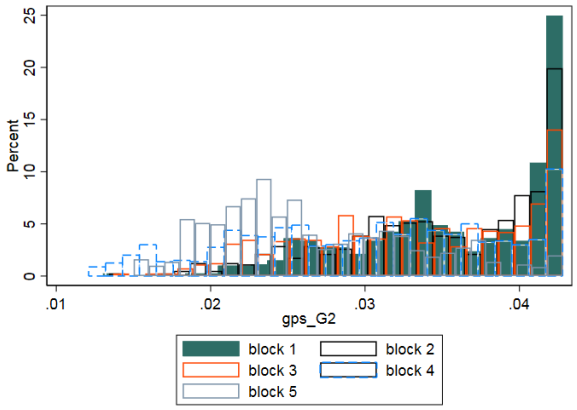
(o) GPS at p90 - after step 3

Sources: RDC of the Federal Statistical Office and Statistical Offices of the Länder (2018a; 2018b), SOEP as well as further data sources as discussed in Chapter 2, own calculations.

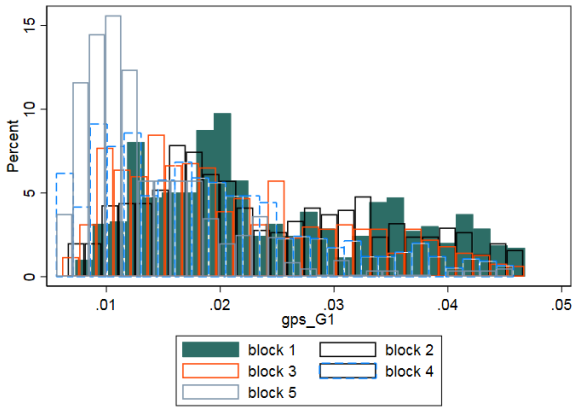
Figure E.2: Common support - A3 sample (46-65 years)



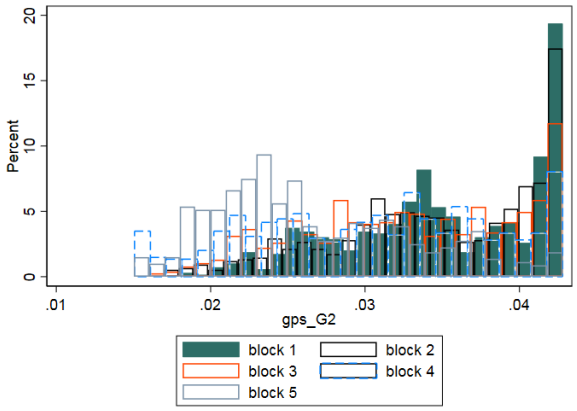
(a) GPS at p10 - before CS condition



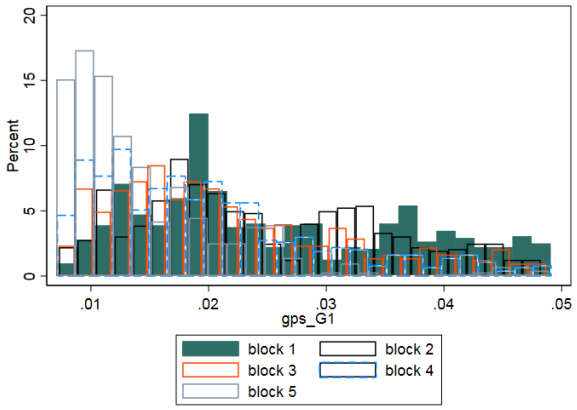
(b) GPS at p30 - before CS condition



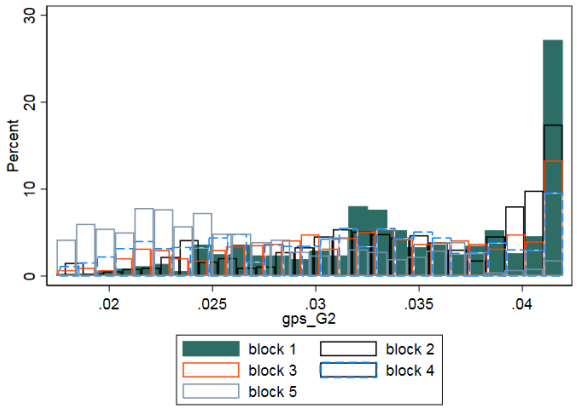
(c) GPS at p10 - after step 1



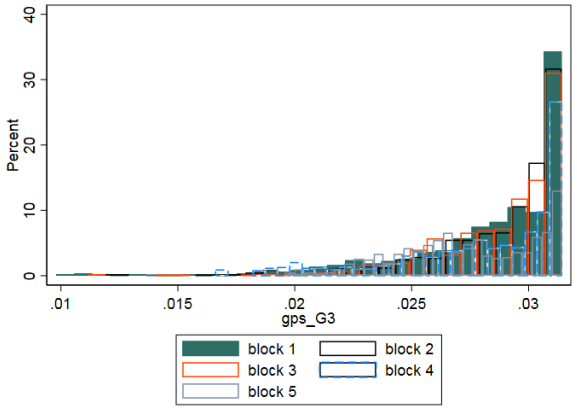
(d) GPS at p30 - after step 1



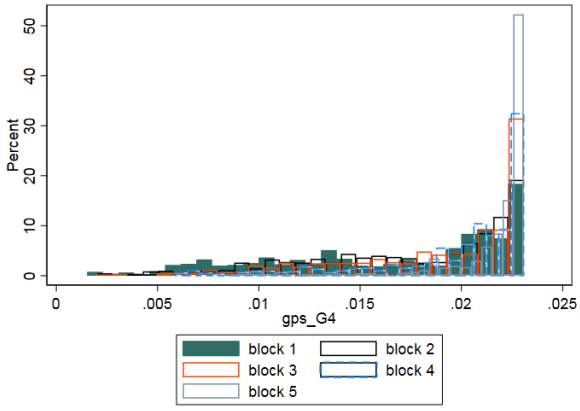
(e) GPS at p10 - after step 3



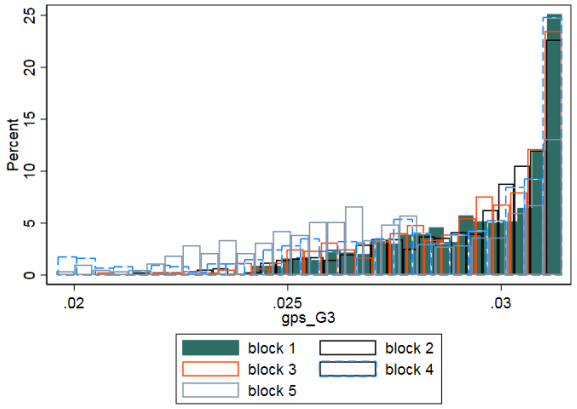
(f) GPS at p30 - after step 3



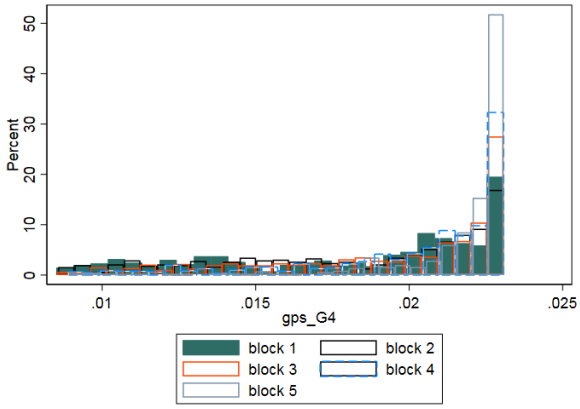
(g) GPS at p50 - before CS condition



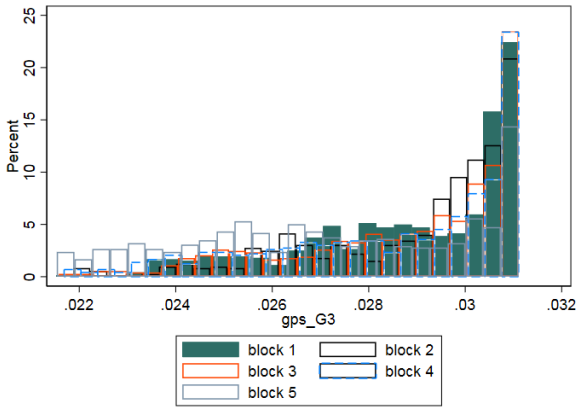
(h) GPS at p70 - before CS condition



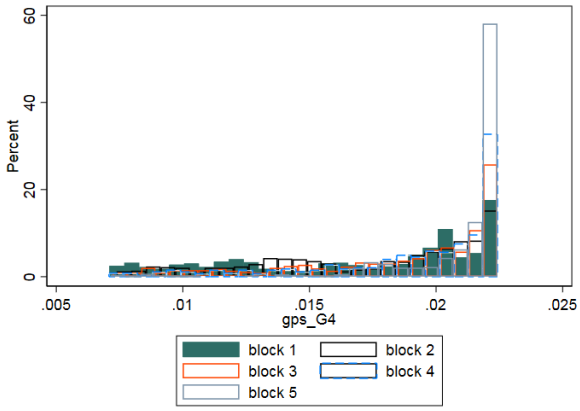
(i) GPS at p50 - after step 1



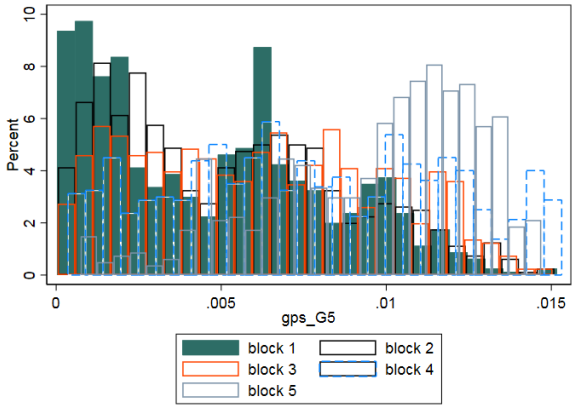
(j) GPS at p70 - after step 1



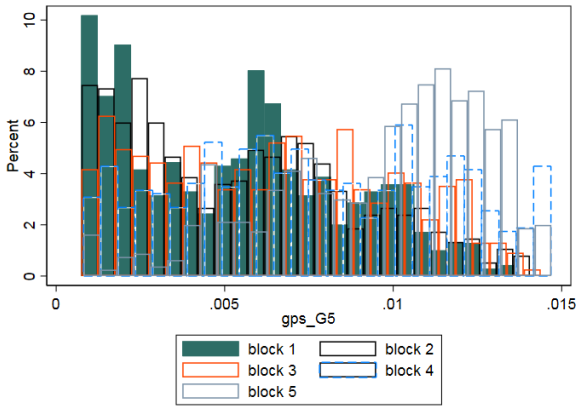
(k) GPS at p50 - after step 3



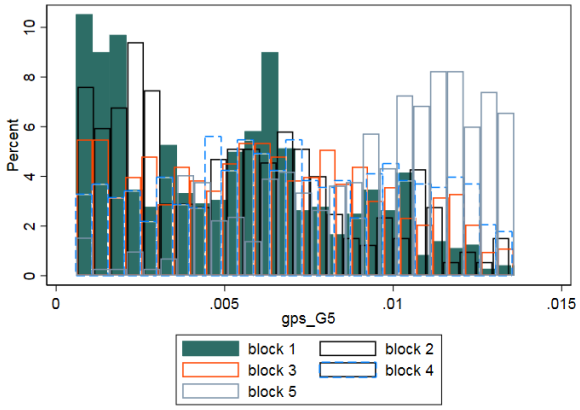
(l) GPS at p70 - after step 3



(m) GPS at p90 - before CS condition



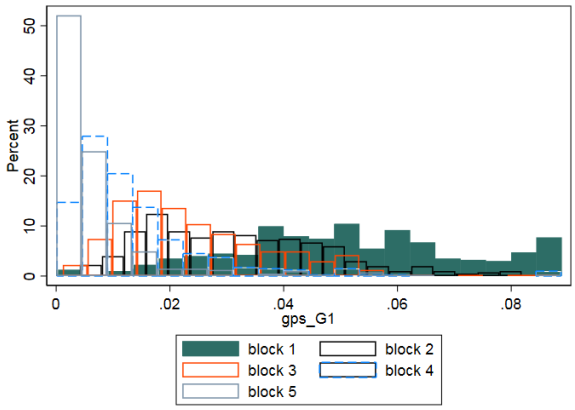
(n) GPS at p90 - after step 1



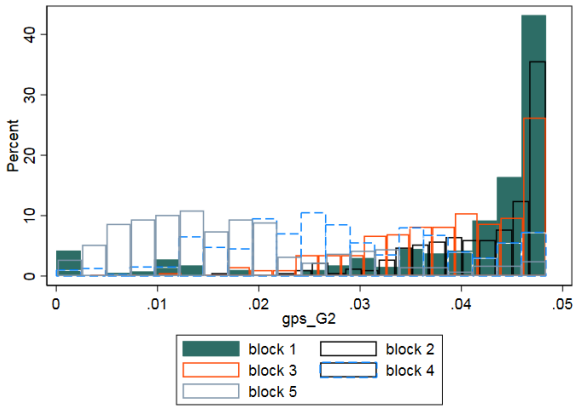
(o) GPS at p90 - after step 3

Sources: RDC of the Federal Statistical Office and Statistical Offices of the Länder (2018a; 2018b), SOEP as well as further data sources as discussed in Chapter 2, own calculations.

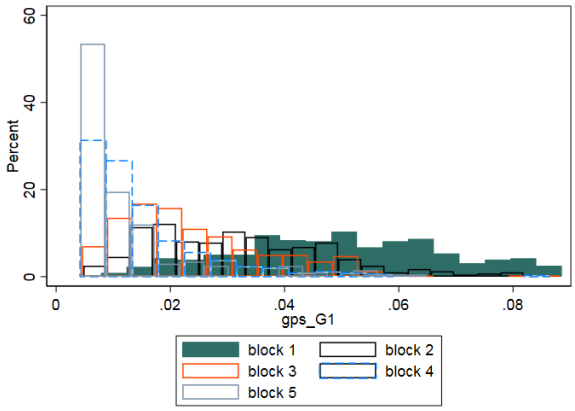
Figure E.3: Common support - A1 with lagged outcome and treatment (17-30 years)



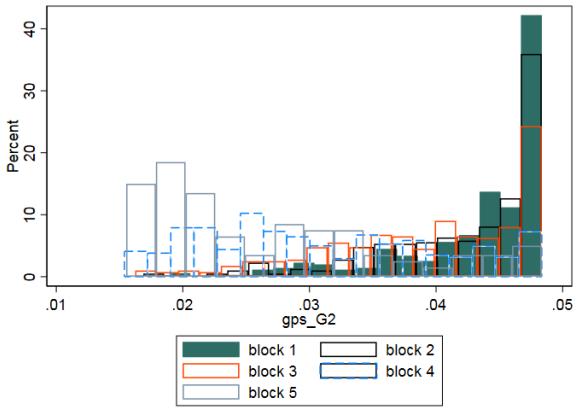
(a) GPS at p10 - before CS condition



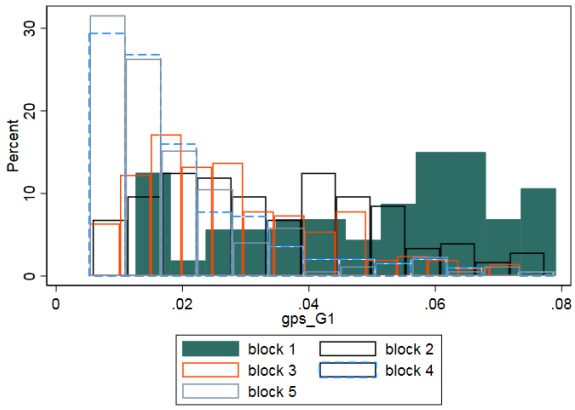
(b) GPS at p30 - before CS condition



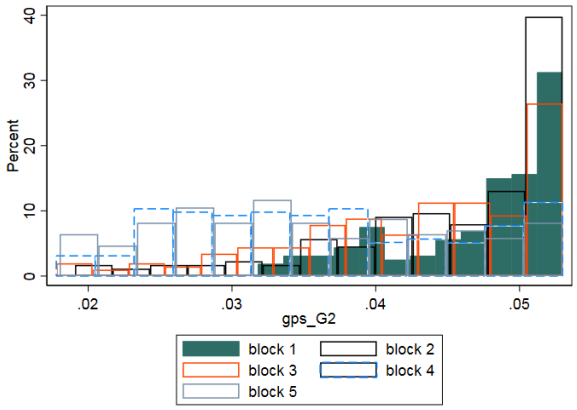
(c) GPS at p10 - after step 1



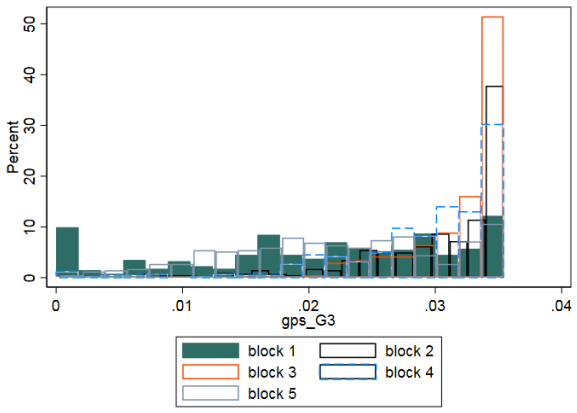
(d) GPS at p30 - after step 1



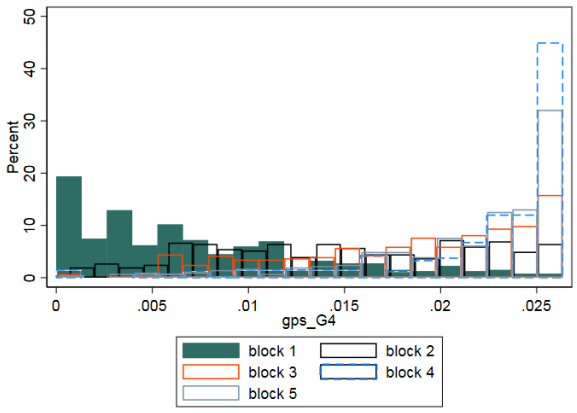
(e) GPS at p10 - after step 3



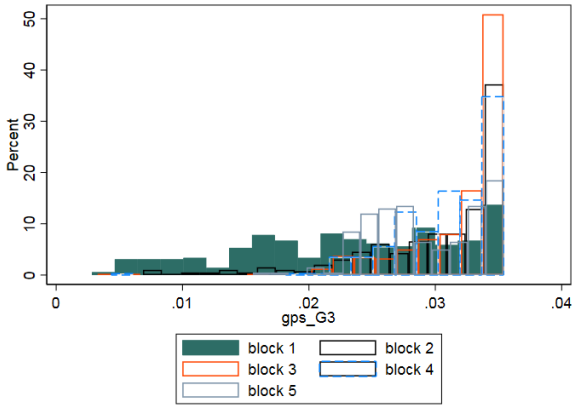
(f) GPS at p30 - after step 3



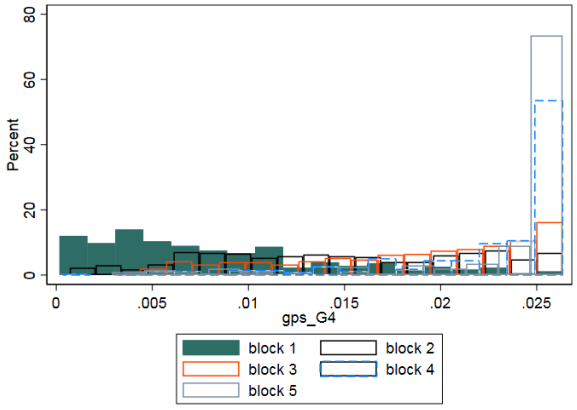
(g) GPS at p50 - before CS condition



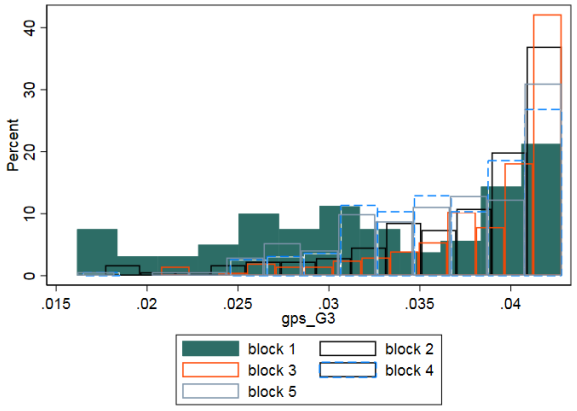
(h) GPS at p70 - before CS condition



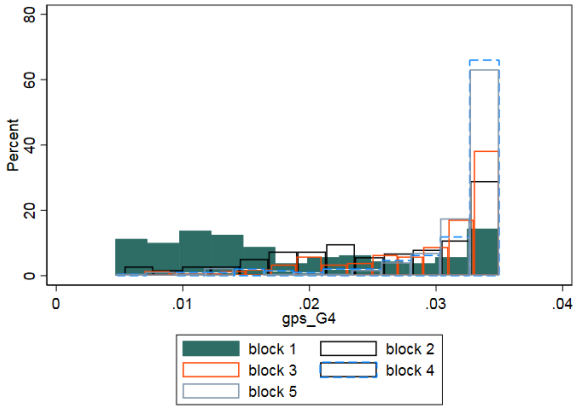
(i) GPS at p50 - after step 1



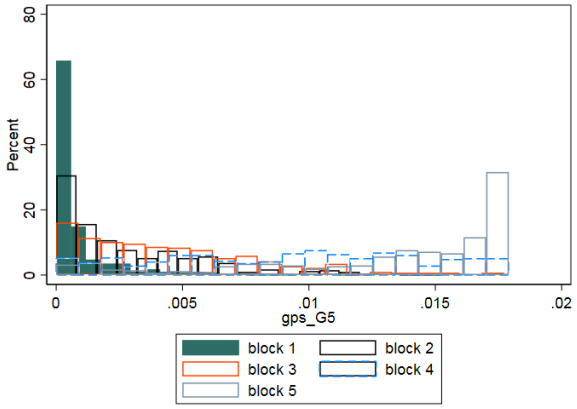
(j) GPS at p70 - after step 1



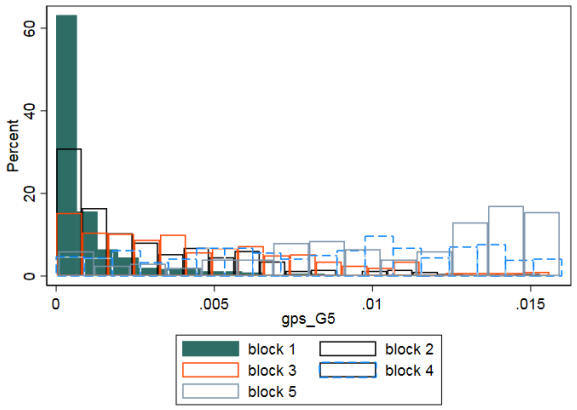
(k) GPS at p50 - after step 3



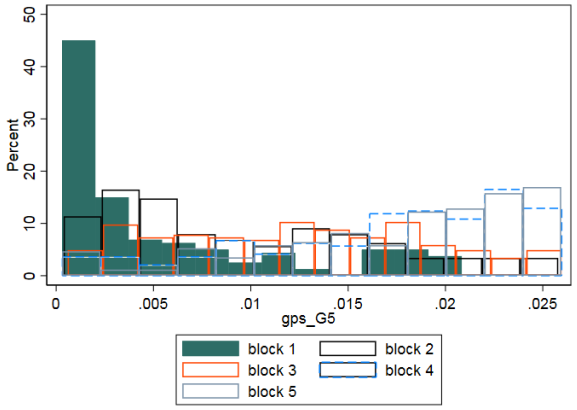
(l) GPS at p70 - after step 3



(m) GPS at p90 - before CS condition



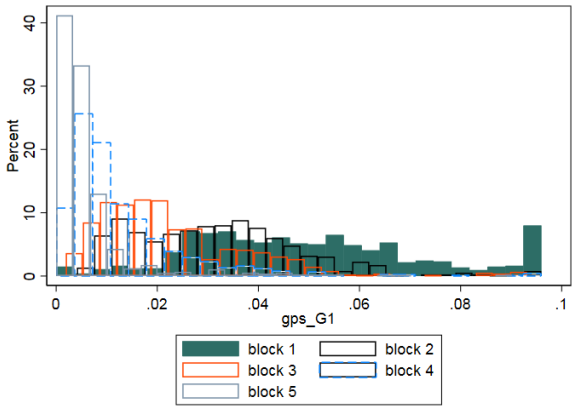
(n) GPS at p90 - after step 1



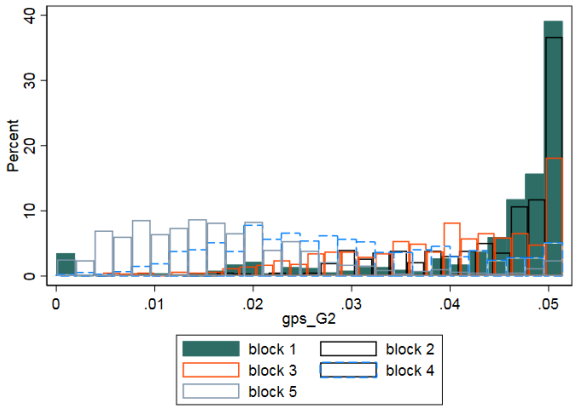
(o) GPS at p90 - after step 3

Sources: RDC of the Federal Statistical Office and Statistical Offices of the Länder (2018a; 2018b), SOEP as well as further data sources as discussed in Chapter 2, own calculations.

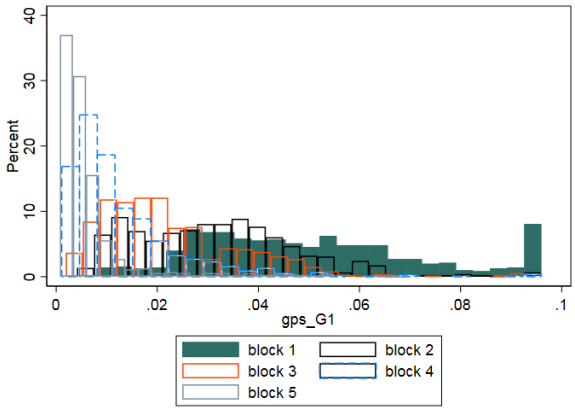
Figure E.4: Common support - A2 with lagged outcome and treatment (31-45 years)



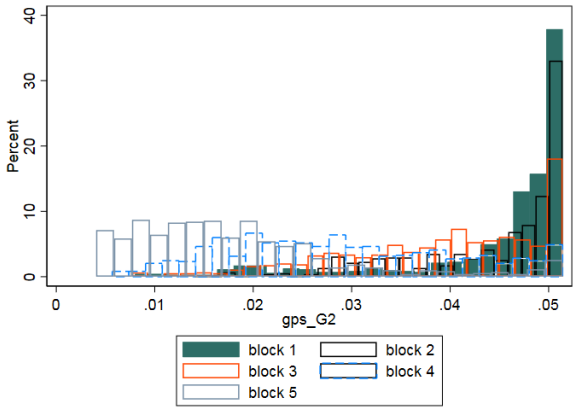
(a) GPS at p10 - before CS condition



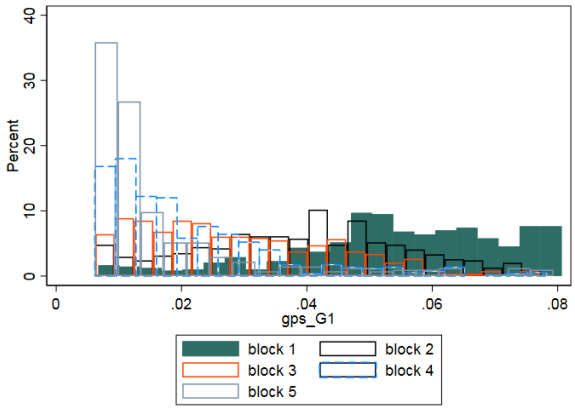
(b) GPS at p30 - before CS condition



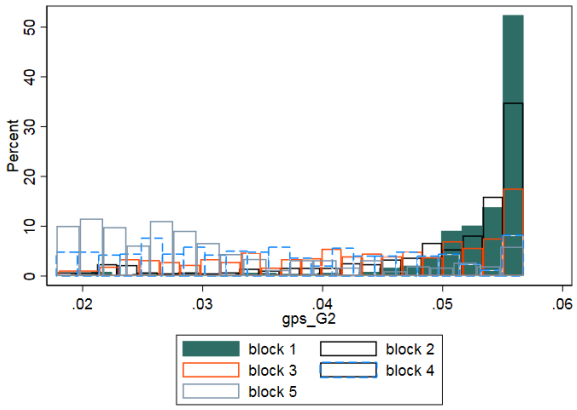
(c) GPS at p10 - after step 1



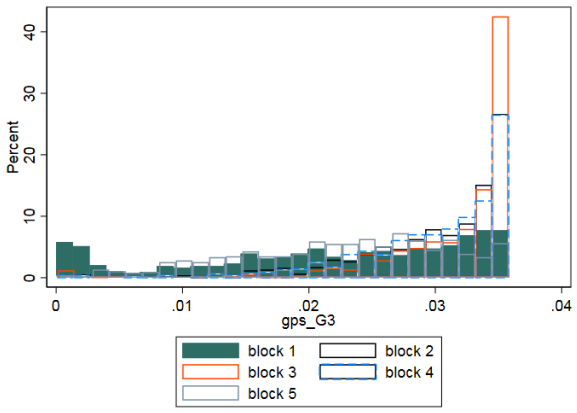
(d) GPS at p30 - after step 1



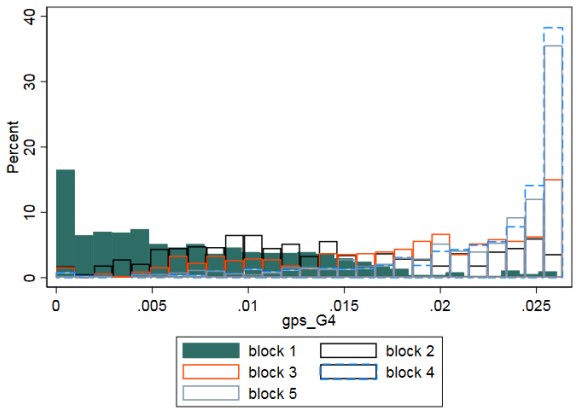
(e) GPS at p10 - after step 3



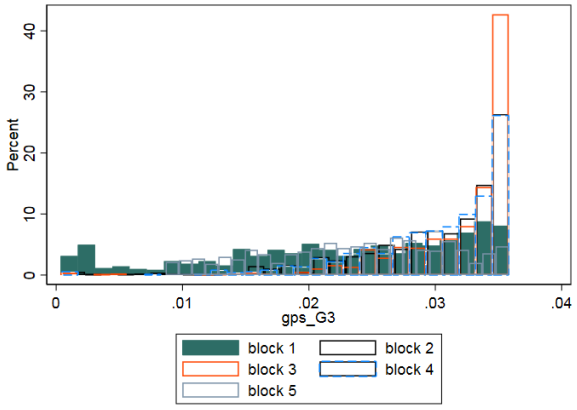
(f) GPS at p30 - after step 3



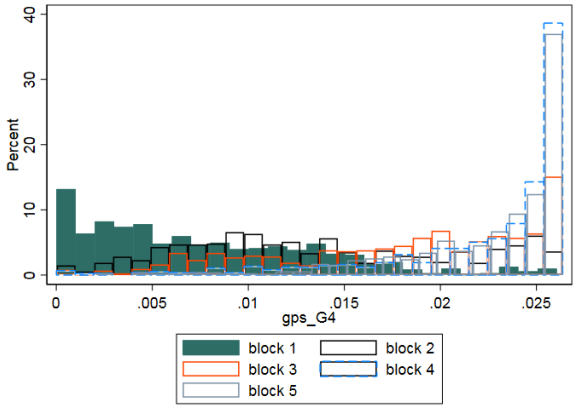
(g) GPS at p50 - before CS condition



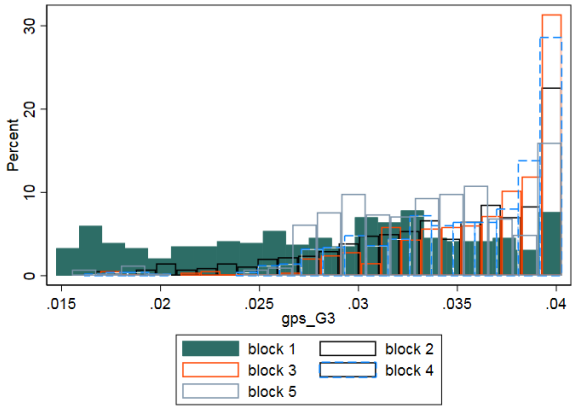
(h) GPS at p70 - before CS condition



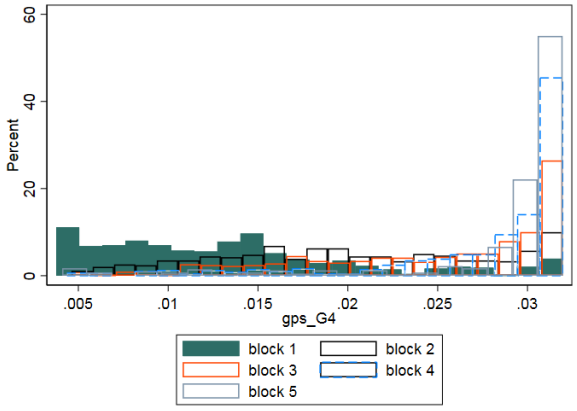
(i) GPS at p50 - after step 1



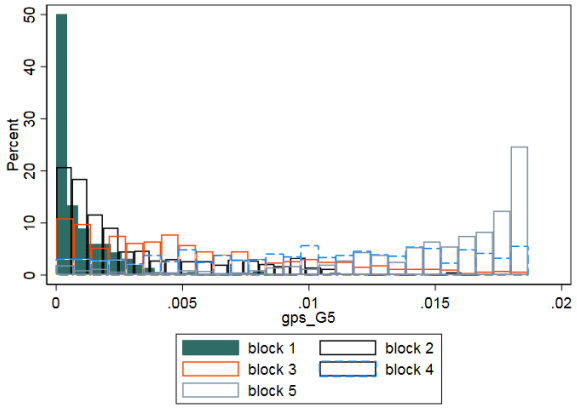
(j) GPS at p70 - after step 1



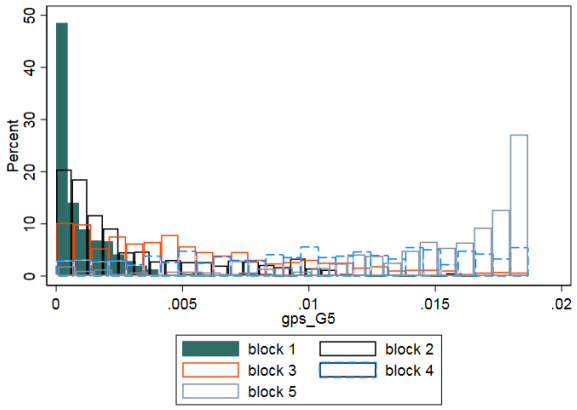
(k) GPS at p50 - after step 3



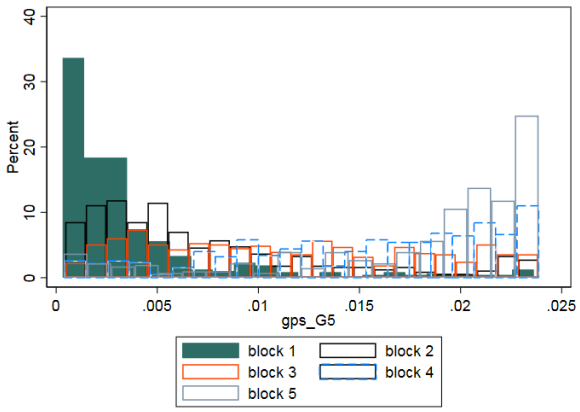
(l) GPS at p70 - after step 3



(m) GPS at p90 - before CS condition



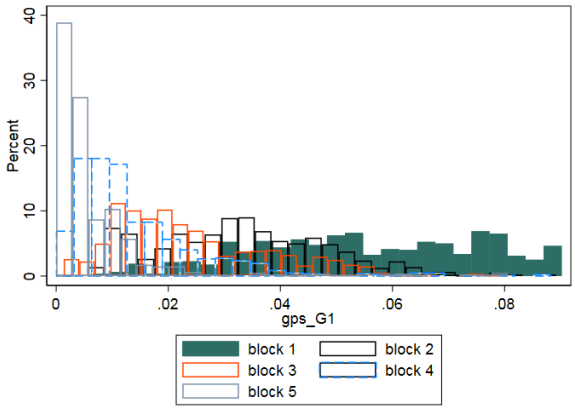
(n) GPS at p90 - after step 1



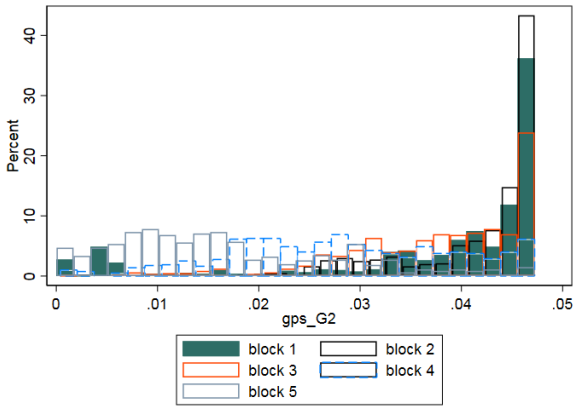
(o) GPS at p90 - after step 3

Sources: RDC of the Federal Statistical Office and Statistical Offices of the Länder (2018a; 2018b), SOEP as well as further data sources as discussed in Chapter 2, own calculations.

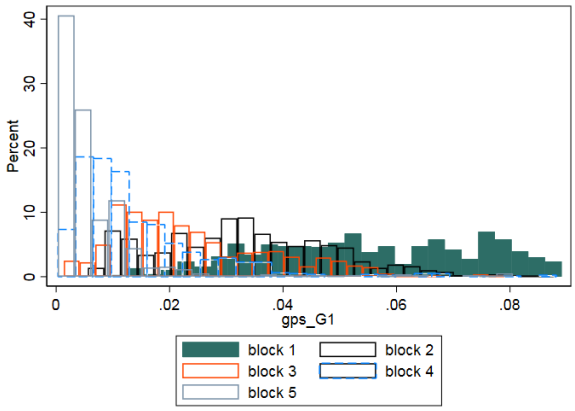
Figure E.5: Common support - A3 with lagged outcome and treatment (46-65 years)



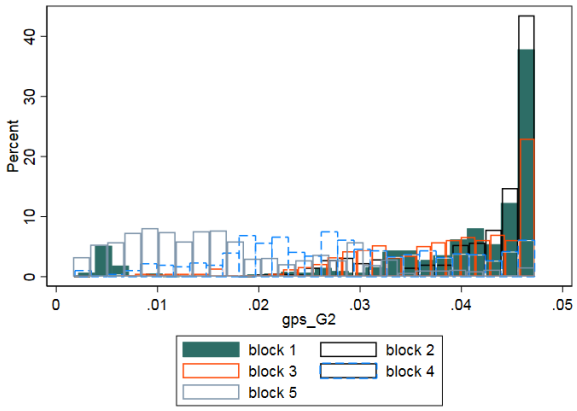
(a) GPS at p10 - before CS condition



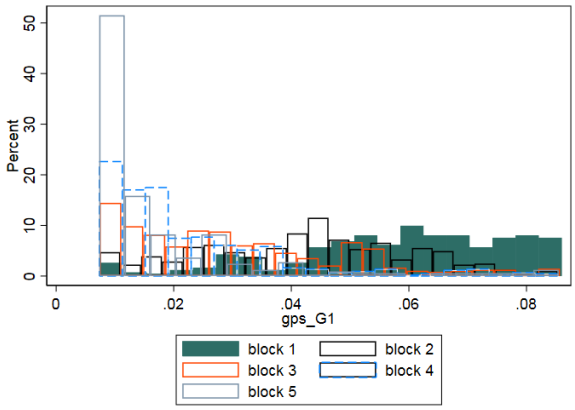
(b) GPS at p30 - before CS condition



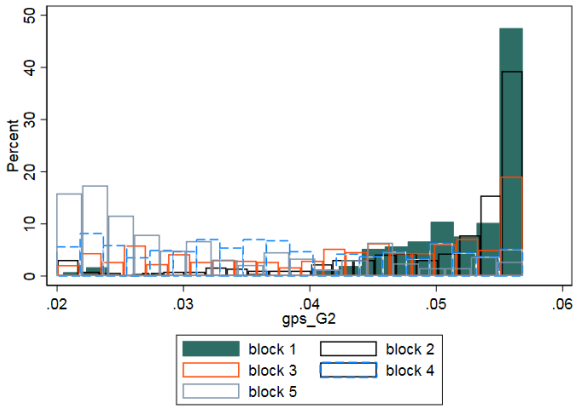
(c) GPS at p10 - after step 1



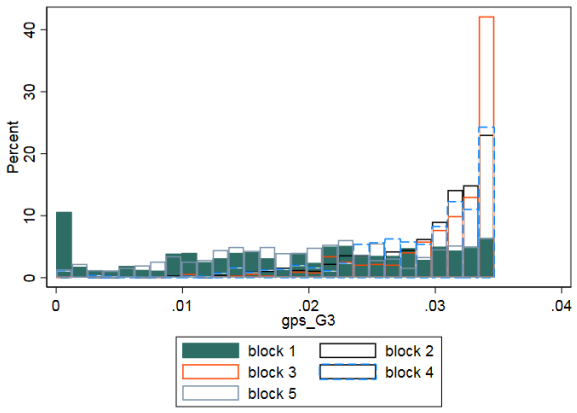
(d) GPS at p30 - after step 1



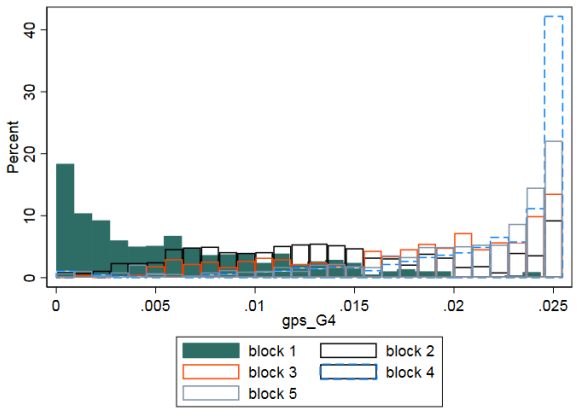
(e) GPS at p10 - after step 3



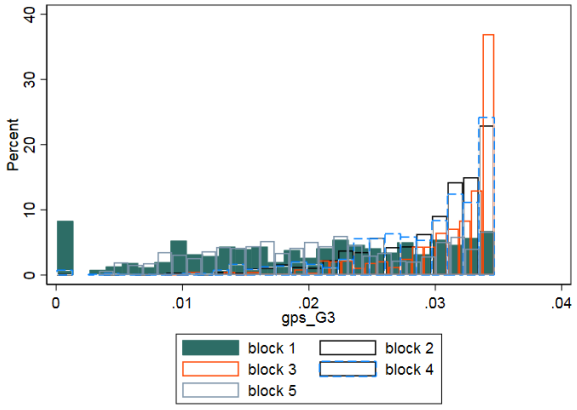
(f) GPS at p30 - after step 3



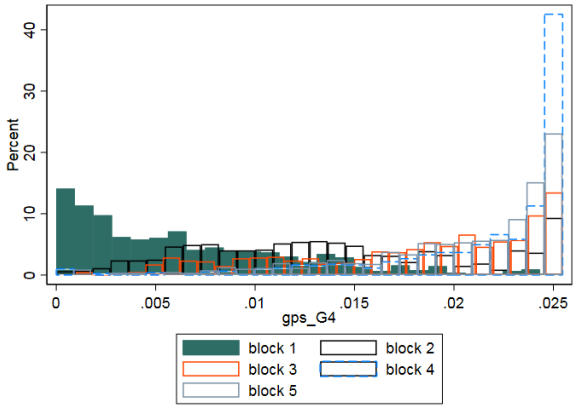
(g) GPS at p50 - before CS condition



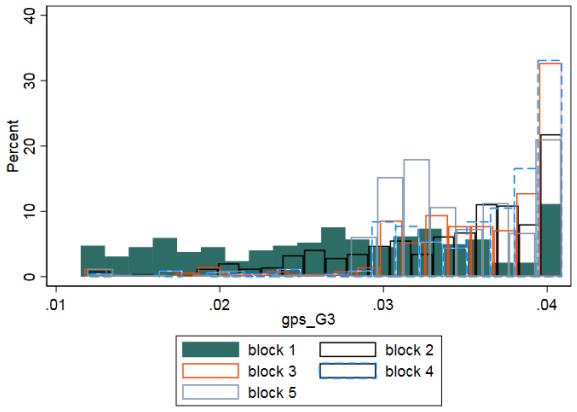
(h) GPS at p70 - before CS condition



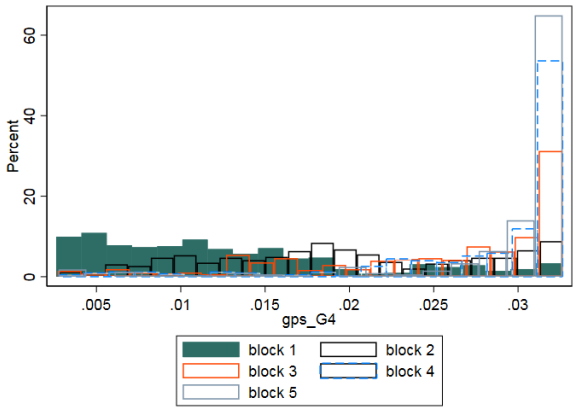
(i) GPS at p50 - after step 1



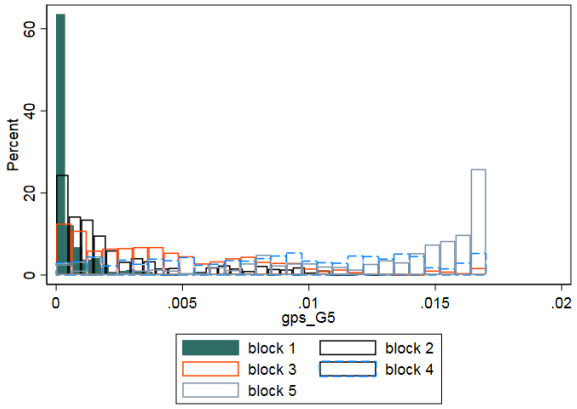
(j) GPS at p70 - after step 1



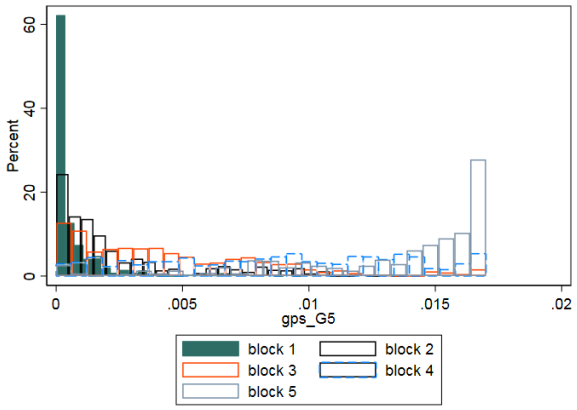
(k) GPS at p50 - after step 3



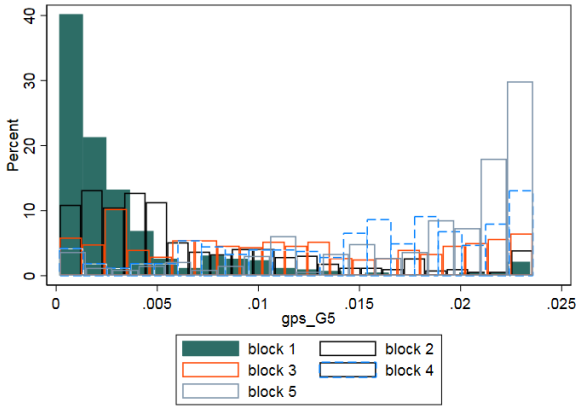
(l) GPS at p70 - after step 3



(m) GPS at p90 - before CS condition



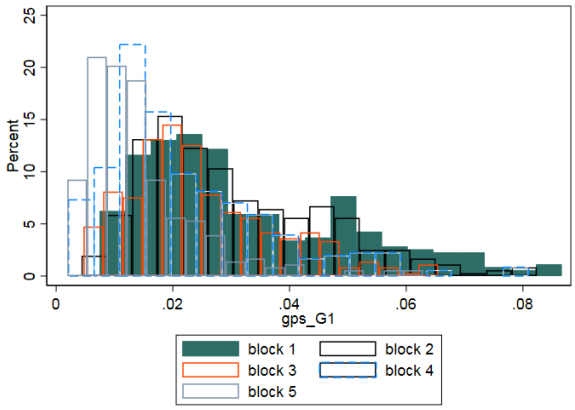
(n) GPS at p90 - after step 1



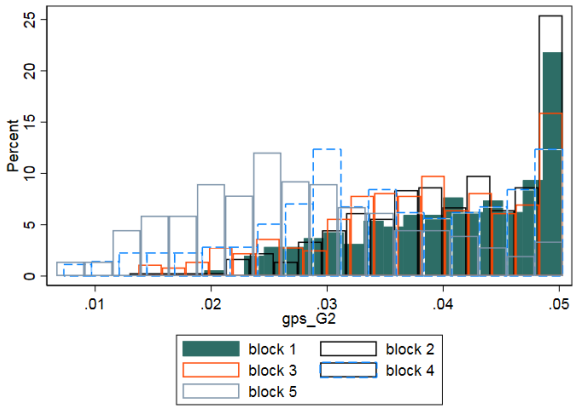
(o) GPS at p90 - after step 3

Sources: RDC of the Federal Statistical Office and Statistical Offices of the Länder (2018a; 2018b), SOEP as well as further data sources as discussed in Chapter 2, own calculations.

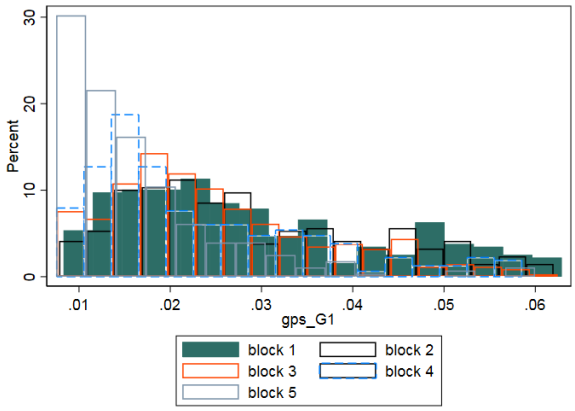
Figure E.6: Common support - A2 sample - small municipalities



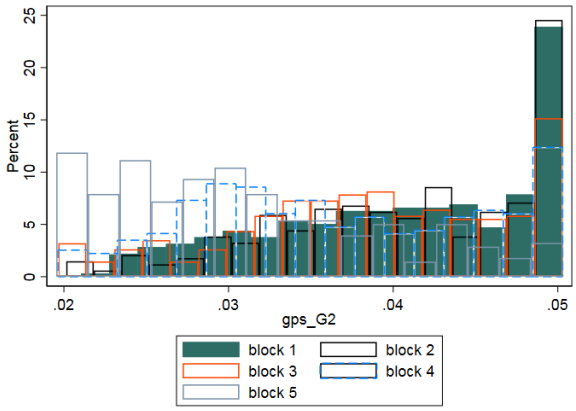
(a) GPS at p10 - before CS condition



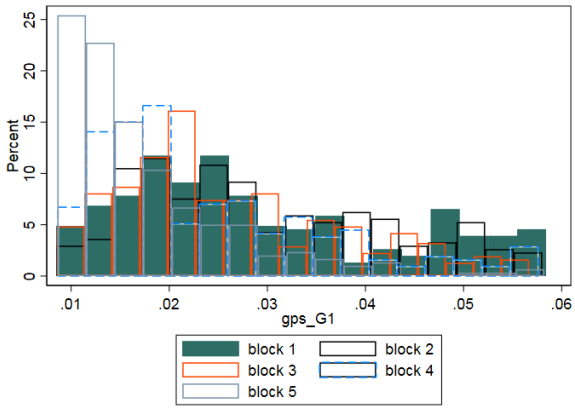
(b) GPS at p30 - before CS condition



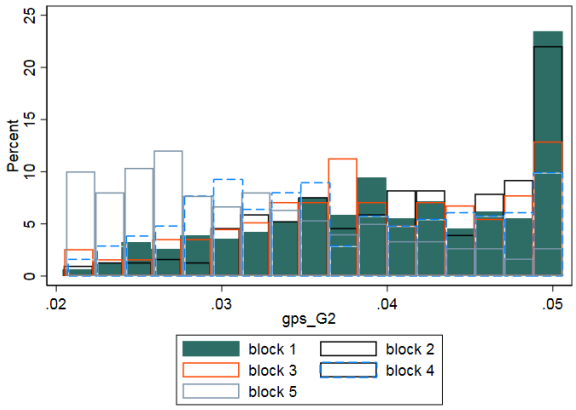
(c) GPS at p10 - after step 1



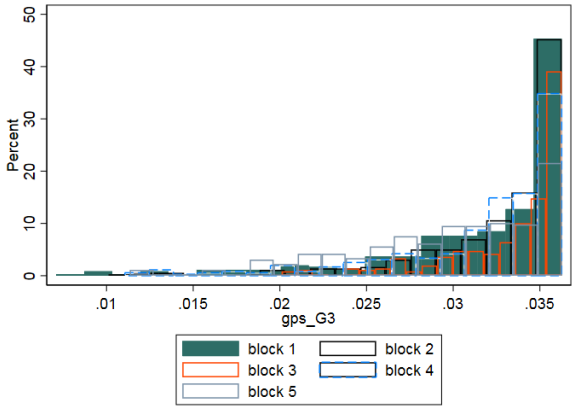
(d) GPS at p30 - after step 1



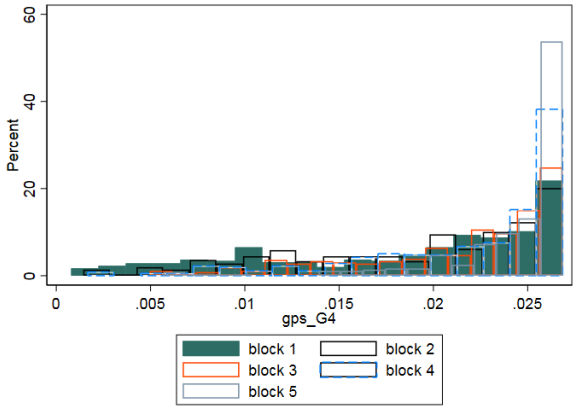
(e) GPS at p10 - after step 3



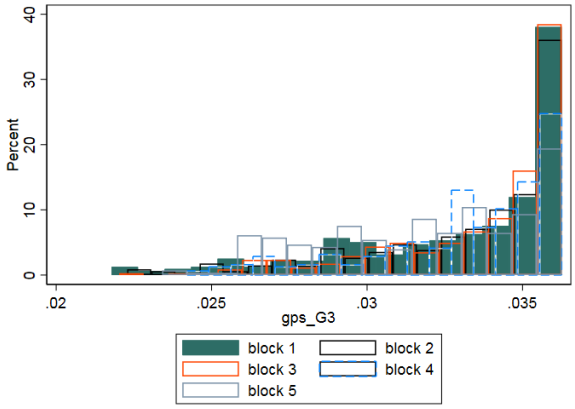
(f) GPS at p30 - after step 3



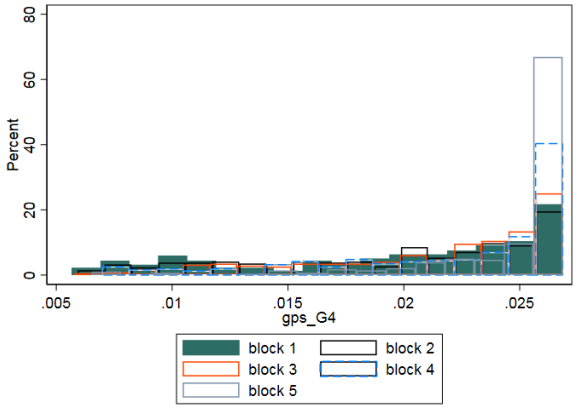
(g) GPS at p50 - before CS condition



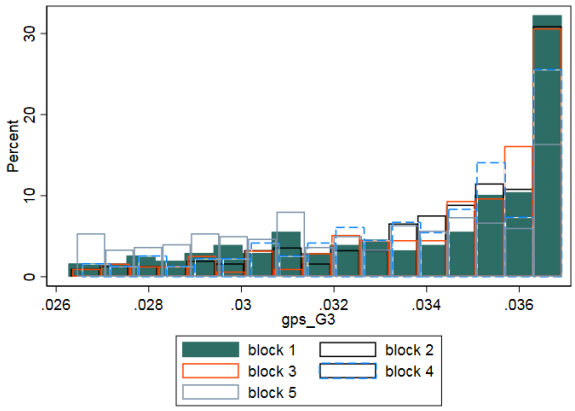
(h) GPS at p70 - before CS condition



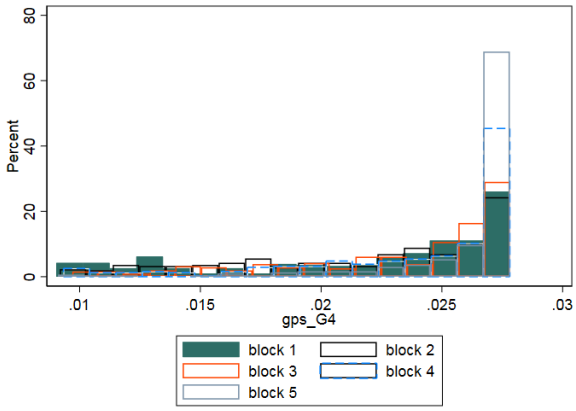
(i) GPS at p50 - after step 1



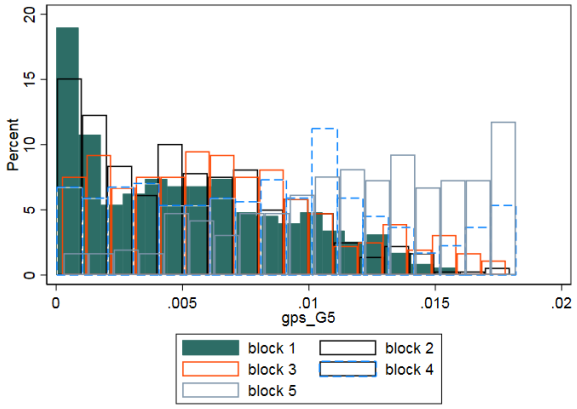
(j) GPS at p70 - after step 1



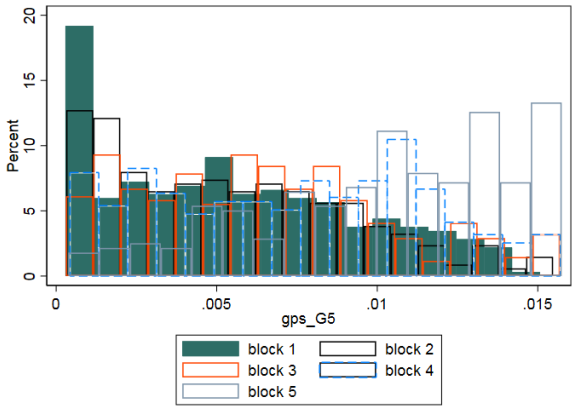
(k) GPS at p50 - after step 3



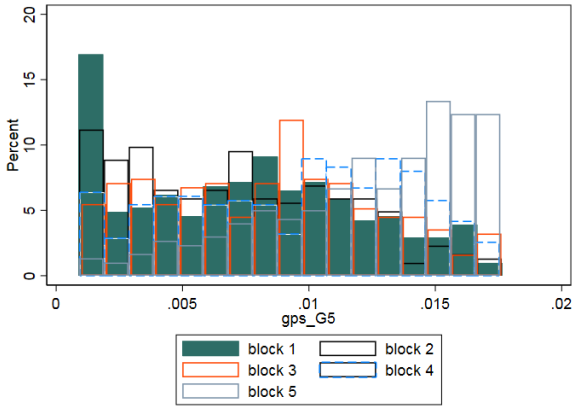
(l) GPS at p70 - after step 3



(m) GPS at p90 - before CS condition



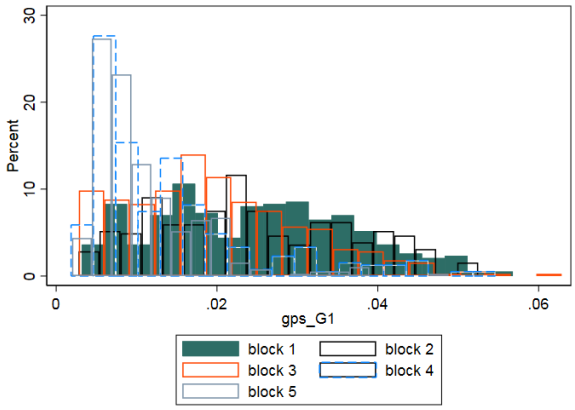
(n) GPS at p90 - after step 1



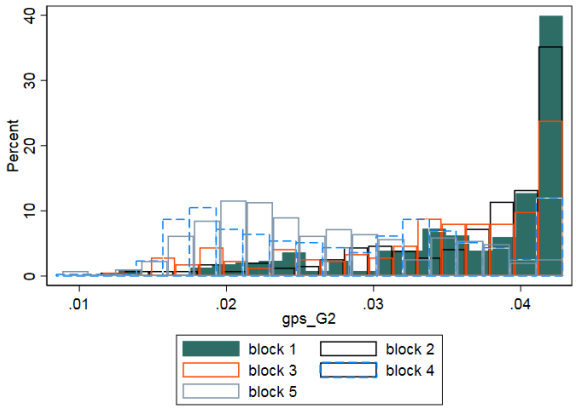
(o) GPS at p90 - after step 3

Sources: RDC of the Federal Statistical Office and Statistical Offices of the Länder (2018a; 2018b), SOEP as well as further data sources as discussed in Chapter 2, own calculations.

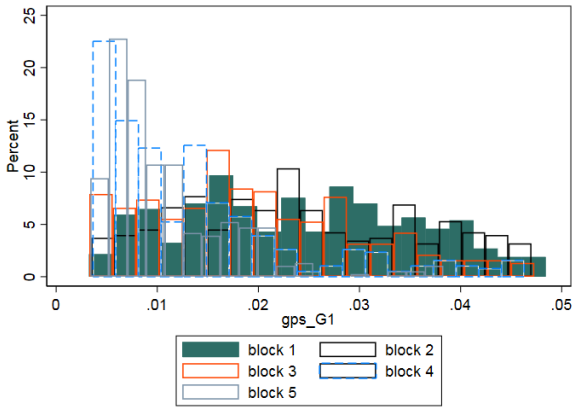
Figure E.7: Common support - A2 sample - large municipalities



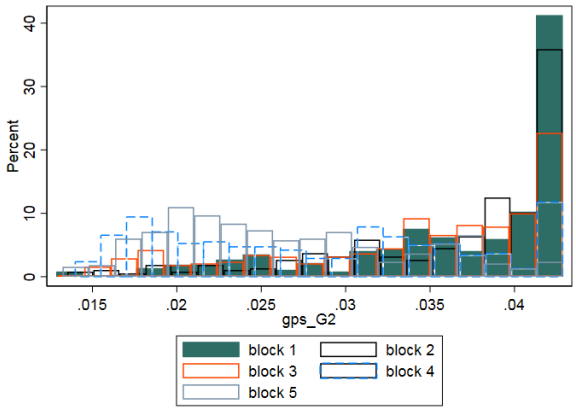
(a) GPS at p10 - before CS condition



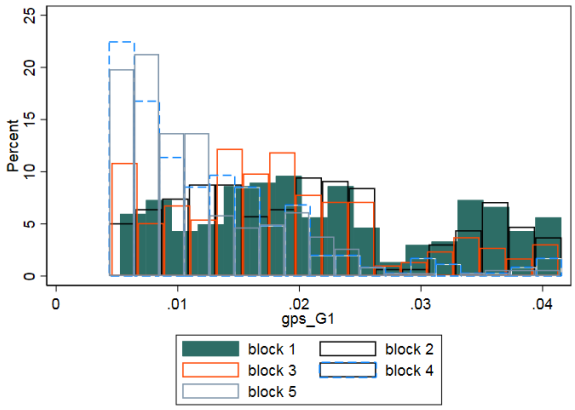
(b) GPS at p30 - before CS condition



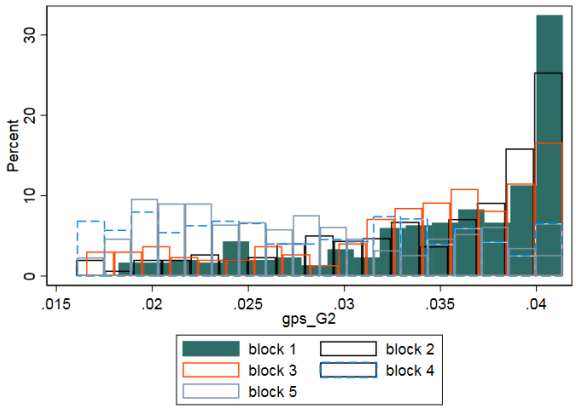
(c) GPS at p10 - after step 1



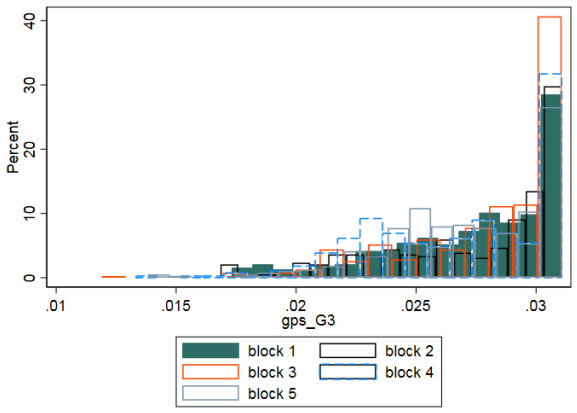
(d) GPS at p30 - after step 1



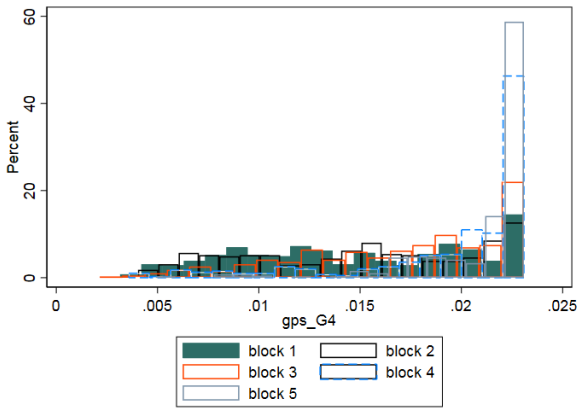
(e) GPS at p10 - after step 3



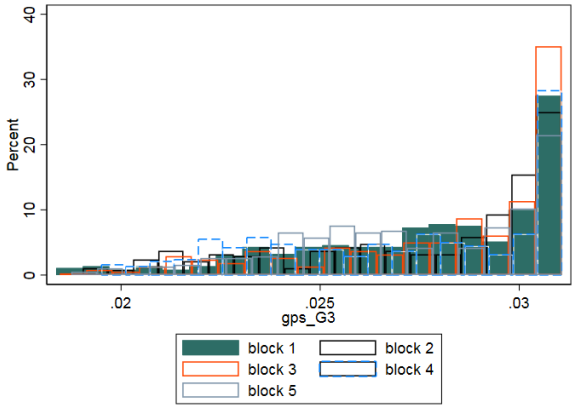
(f) GPS at p30 - after step 3



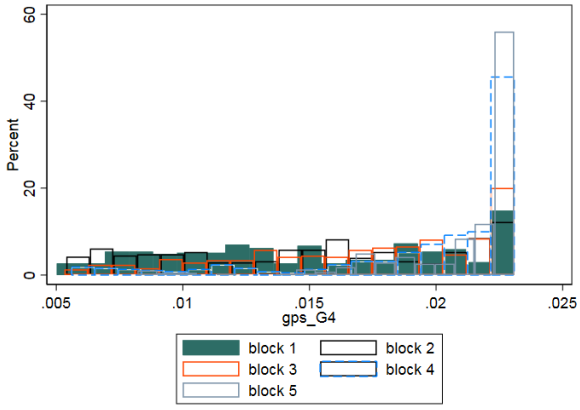
(g) GPS at p50 - before CS condition



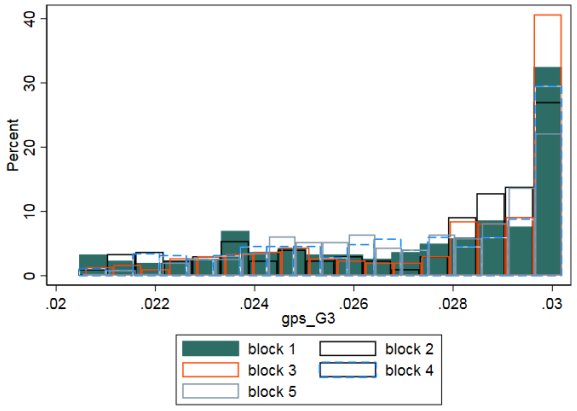
(h) GPS at p70 - before CS condition



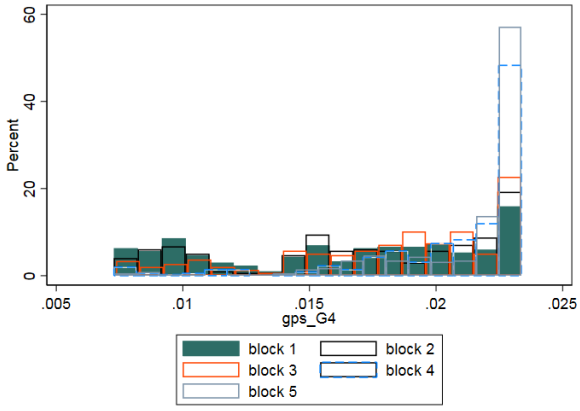
(i) GPS at p50 - after step 1



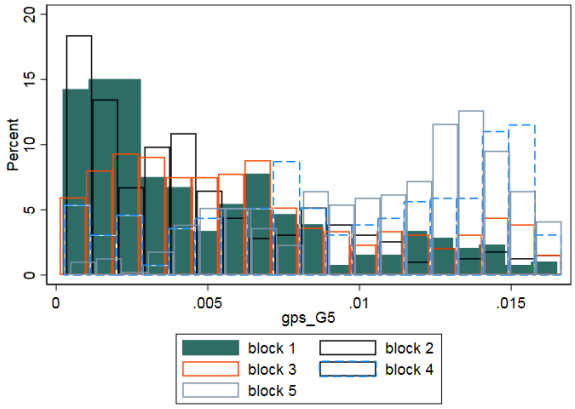
(j) GPS at p70 - after step 1



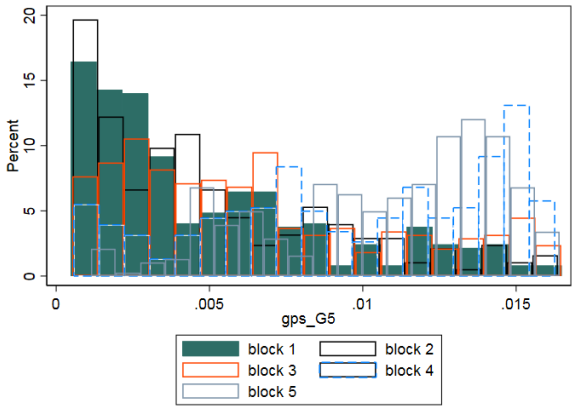
(k) GPS at p50 - after step 3



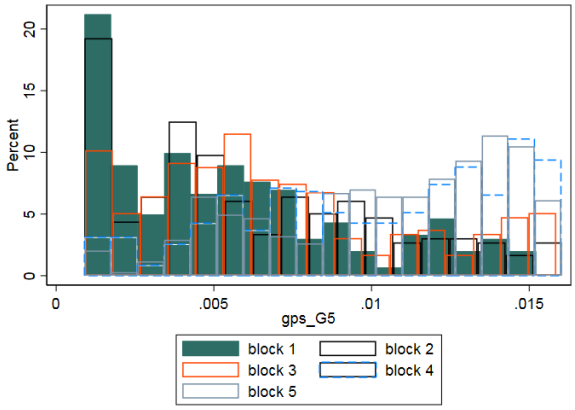
(l) GPS at p70 - after step 3



(m) GPS at p90 - before CS condition



(n) GPS at p90 - after step 1



(o) GPS at p90 - after step 3

Sources: RDC of the Federal Statistical Office and Statistical Offices of the Länder (2018a; 2018b), SOEP as well as further data sources as discussed in Chapter 2, own calculations.

Table E.8: Balancing before GPS adjustment - A1 sample (17-30 years of age)

	Block 1 vs 2			Block 1 vs 3			Block 1 vs 4			Block 1 vs 5			Block 2 vs 3		
	Mean 1	Mean 2	% bias	Mean 1	Mean 3	% bias	Mean 1	Mean 4	% bias	Mean 1	Mean 5	% bias	Mean 2	Mean 3	% bias
ln(Population per km2/100)	1.32	1.20	10.05	1.32	1.57	-20.03	1.32	1.66	-26.05	1.32	1.87	-47.63	1.20	1.57	-31.43
ln(Average rain/100)	2.00	1.97	13.63	2.00	2.03	-17.84	2.00	2.05	-24.71	2.00	2.12	-60.32	1.97	2.03	-31.19
ln(Proportion recreational area pc in sqm/100)	-0.73	-0.76	3.20	-0.73	-0.84	14.86	-0.73	-0.96	30.69	-0.73	-1.15	60.21	-0.76	-0.84	10.30
Unemployment mean	9.77	9.73	0.77	9.77	8.63	28.44	9.77	8.16	40.21	9.77	6.92	75.91	9.73	8.63	27.01
Commuters mean	-34.64	-43.69	10.91	-34.64	-33.08	-1.83	-34.64	-21.41	-18.27	-34.64	-16.86	-24.04	-43.69	-33.08	-11.95
Age	24.19	24.09	2.60	24.19	23.93	6.37	24.19	24.28	-1.99	24.19	24.15	1.05	24.09	23.93	3.82
Number of children	0.70	0.55	16.20	0.70	0.67	3.35	0.70	0.67	3.32	0.70	0.56	15.27	0.55	0.67	-12.22
Male	0.45	0.44	0.01	0.45	0.46	-0.01	0.45	0.43	0.02	0.45	0.48	-0.03	0.44	0.46	-0.02
Married	0.23	0.20	0.02	0.23	0.19	0.04	0.23	0.30	-0.07	0.23	0.24	-0.01	0.20	0.19	0.01
Divorced	0.01	0.01	0.00	0.01	0.01	0.00	0.01	0.02	-0.01	0.01	0.01	0.00	0.01	0.01	0.00
German	0.92	0.96	-0.04	0.92	0.91	0.01	0.92	0.85	0.07	0.92	0.84	0.08	0.96	0.91	0.05
Other graduation diploma	0.03	0.03	0.00	0.03	0.03	0.00	0.03	0.04	-0.01	0.03	0.04	-0.01	0.03	0.03	0.00
Intermediate secondary school	0.36	0.42	-0.06	0.36	0.38	-0.02	0.36	0.32	0.03	0.36	0.28	0.08	0.42	0.38	0.03
Upper secondary school	0.26	0.27	-0.01	0.26	0.27	-0.02	0.26	0.27	-0.02	0.26	0.33	-0.08	0.27	0.27	-0.01
Not yet graduated	0.10	0.11	-0.01	0.10	0.13	-0.03	0.10	0.12	-0.02	0.10	0.10	0.00	0.11	0.13	-0.02
In education	0.26	0.26	0.00	0.26	0.24	0.02	0.26	0.20	0.06	0.26	0.30	-0.04	0.26	0.24	0.02
No vocational degree	0.22	0.22	0.00	0.22	0.25	-0.03	0.22	0.28	-0.07	0.22	0.25	-0.04	0.22	0.25	-0.03
Degree below university	0.52	0.54	-0.02	0.52	0.49	0.03	0.52	0.49	0.02	0.52	0.43	0.08	0.54	0.49	0.05
University	0.06	0.05	0.01	0.06	0.09	-0.03	0.06	0.07	-0.01	0.06	0.09	-0.03	0.05	0.09	-0.04
Northern states	0.22	0.14	0.08	0.22	0.15	0.07	0.22	0.05	0.17	0.22	0.02	0.20	0.14	0.15	-0.01
Western/Central states	0.07	0.07	0.00	0.07	0.13	-0.06	0.07	0.16	-0.09	0.07	0.21	-0.14	0.07	0.13	-0.06
Southern states	0.16	0.17	-0.01	0.16	0.16	0.00	0.16	0.35	-0.19	0.16	0.44	-0.28	0.17	0.16	0.01
Eastern states	0.33	0.46	-0.13	0.33	0.29	0.04	0.33	0.26	0.07	0.33	0.12	0.21	0.46	0.29	0.17

	Block 2 vs 4			Block 2 vs 5			Block 3 vs 4			Block 3 vs 5			Block 4 vs 5		
	Mean 2	Mean 4	% bias	Mean 2	Mean 5	% bias	Mean 3	Mean 4	% bias	Mean 3	Mean 5	% bias	Mean 4	Mean 5	% bias
ln(Population per km2/100)	1.20	1.66	-37.43	1.20	1.87	-61.42	1.57	1.66	-6.75	1.57	1.87	-27.71	1.66	1.87	-19.67
ln(Average rain/100)	1.97	2.05	-37.02	1.97	2.12	-72.37	2.03	2.05	-8.26	2.03	2.12	-43.40	2.05	2.12	-32.11
ln(Proportion recreational area pc in sqm/100)	-0.76	-0.96	25.04	-0.76	-1.15	51.58	-0.84	-0.96	17.03	-0.84	-1.15	47.20	-0.96	-1.15	27.97
Unemployment mean	9.73	8.16	38.53	9.73	6.92	73.17	8.63	8.16	12.47	8.63	6.92	48.20	8.16	6.92	34.58
Commuters mean	-43.69	-21.41	-29.17	-43.69	-16.86	-34.45	-33.08	-21.41	-14.77	-33.08	-16.86	-20.17	-21.41	-16.86	-6.84
Age	24.09	24.28	-4.55	24.09	24.15	-1.56	23.93	24.28	-8.22	23.93	24.15	-5.36	24.28	24.15	3.03
Number of children	0.55	0.67	-13.56	0.55	0.56	-1.06	0.67	0.67	-0.26	0.67	0.56	11.31	0.67	0.56	12.57
Male	0.44	0.43	0.01	0.44	0.48	-0.04	0.46	0.43	0.03	0.46	0.48	-0.02	0.43	0.48	-0.05
Married	0.20	0.30	-0.10	0.20	0.24	-0.04	0.19	0.30	-0.11	0.19	0.24	-0.05	0.30	0.24	0.06
Divorced	0.01	0.02	-0.01	0.01	0.01	0.00	0.01	0.02	-0.01	0.01	0.01	0.00	0.02	0.01	0.01
German	0.96	0.85	0.11	0.96	0.84	0.12	0.91	0.85	0.05	0.91	0.84	0.07	0.85	0.84	0.02
Other graduation diploma	0.03	0.04	-0.02	0.03	0.04	-0.02	0.03	0.04	-0.01	0.03	0.04	-0.01	0.04	0.04	0.00
Intermediate secondary school	0.42	0.32	0.09	0.42	0.28	0.14	0.38	0.32	0.06	0.38	0.28	0.10	0.32	0.28	0.05
Upper secondary school	0.27	0.27	-0.01	0.27	0.33	-0.06	0.27	0.27	0.00	0.27	0.33	-0.06	0.27	0.33	-0.06
Not yet graduated	0.11	0.12	-0.01	0.11	0.10	0.01	0.13	0.12	0.01	0.13	0.10	0.03	0.12	0.10	0.02
In education	0.26	0.20	0.06	0.26	0.30	-0.04	0.24	0.20	0.04	0.24	0.30	-0.06	0.20	0.30	-0.10
No vocational degree	0.22	0.28	-0.07	0.22	0.25	-0.04	0.25	0.28	-0.03	0.25	0.25	0.00	0.28	0.25	0.03
Degree below university	0.54	0.49	0.05	0.54	0.43	0.10	0.49	0.49	0.00	0.49	0.43	0.05	0.49	0.43	0.06
University	0.05	0.07	-0.02	0.05	0.09	-0.04	0.09	0.07	0.02	0.09	0.09	0.00	0.07	0.09	-0.02
Northern states	0.14	0.05	0.08	0.14	0.02	0.12	0.15	0.05	0.10	0.15	0.02	0.13	0.05	0.02	0.04
Western/Northern states	0.07	0.16	-0.08	0.07	0.21	-0.14	0.13	0.16	-0.02	0.13	0.21	-0.08	0.16	0.21	-0.05
Southern states	0.17	0.35	-0.18	0.17	0.44	-0.27	0.16	0.35	-0.19	0.16	0.44	-0.28	0.35	0.44	-0.09
Eastern states	0.46	0.26	0.20	0.46	0.12	0.34	0.29	0.26	0.03	0.29	0.12	0.17	0.26	0.12	0.14

Sources: RDC of the Federal Statistical Office and Statistical Offices of the Länder (2018a; 2018b), SOEP as well as further data sources as discussed in Chapter 2, own calculations.

Table E.9: Balancing after step 2 of common support procedure - A1 sample

	Block 1 vs 2	Block 1 vs 3	Block 1 vs 4	Block 1 vs 5	Block 2 vs 3	Block 2 vs 4	Block 2 vs 5	Block 3 vs 4	Block 3 vs 5	Block 4 vs 5
ln(Population per km^2/100)	16.12	0.18	8.92	-1.39	-16.83	-7.42	-20.02	9.22	-1.70	-11.71
ln(Average rain/100)	10.63	3.67	19.78	2.76	-7.23	9.25	-7.98	16.61	-0.87	-17.24
ln(Proportion recreational area pc in sqm/100)	8.81	2.23	5.27	3.03	-7.08	-4.01	-6.51	3.24	0.79	-2.55
Unemployment mean	5.62	9.74	-4.22	-4.35	3.67	-9.96	-10.46	-14.59	-15.54	0.16
Commuters mean	11.23	12.92	5.85	23.00	1.93	-6.65	9.88	-8.60	7.43	19.21
Age	-0.03	2.69	0.26	-0.36	2.72	0.29	-0.32	-2.38	-3.05	-0.61
Number of children	16.29	3.53	-0.46	11.16	-12.11	-17.93	-5.89	-4.18	7.06	12.44
Male	0.00	0.01	0.02	-0.02	0.01	0.02	-0.02	0.01	-0.03	-0.04
Married	0.03	0.05	-0.04	0.03	0.02	-0.07	0.00	-0.09	-0.02	0.07
Divorced	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
German	-0.02	0.01	0.03	0.02	0.03	0.05	0.03	0.02	0.00	-0.02
Other graduation diploma	-0.01	0.00	-0.02	0.00	0.01	0.00	0.01	-0.02	0.00	0.01
Intermediate secondary school	-0.07	-0.07	-0.05	-0.06	0.00	0.02	0.01	0.02	0.02	0.00
Upper secondary school	0.04	0.03	0.06	-0.01	-0.01	0.02	-0.05	0.03	-0.04	-0.07
Not yet graduated	0.00	-0.01	-0.01	0.04	-0.02	-0.02	0.03	0.00	0.05	0.05
In education	0.04	0.07	0.12	-0.03	0.03	0.07	-0.08	0.04	-0.11	-0.15
No vocational degree	0.00	-0.03	-0.05	0.02	-0.03	-0.05	0.02	-0.02	0.05	0.07
Degree below university	-0.07	-0.04	-0.06	-0.02	0.03	0.01	0.05	-0.02	0.02	0.04
University	0.02	-0.02	0.01	0.00	-0.04	0.00	-0.02	0.03	0.02	-0.01
Northern states	0.05	-0.01	0.05	0.06	-0.06	0.00	0.01	0.06	0.07	0.01
Western states	0.01	-0.02	0.00	0.04	-0.03	-0.01	0.03	0.02	0.06	0.04
Southern states	-0.01	0.06	-0.01	0.01	0.07	0.00	0.02	-0.07	-0.05	0.02
Eastern states	-0.11	-0.04	-0.11	-0.08	0.07	0.01	0.03	-0.07	-0.04	0.02

Sources: RDC of the Federal Statistical Office and Statistical Offices of the Länder (2018a; 2018b), SOEP as well as further data sources as discussed in Chapter 2, own calculations.

Table E.10: Balancing before GPS adjustment - A3 sample (46-65 years of age)

	Block 1 vs 2			Block 1 vs 3			Block 1 vs 4			Block 1 vs 5			Block 2 vs 3		
	Mean 1	Mean 2	% bias	Mean 1	Mean 3	% bias	Mean 1	Mean 4	% bias	Mean 1	Mean 5	% bias	Mean 2	Mean 3	% bias
ln(Population per km2/100)	1.25	1.19	4.46	1.25	1.35	-8.48	1.25	1.66	-32.86	1.25	1.88	-55.12	1.19	1.35	-13.91
ln(Average rain/100)	2.01	1.99	5.00	2.01	2.04	-16.94	2.01	2.06	-26.16	2.01	2.12	-56.42	1.99	2.04	-22.46
ln(Proportion recreational area pc in sqm/100)	-0.60	-0.80	24.21	-0.60	-0.82	29.25	-0.60	-0.96	47.61	-0.60	-1.14	77.80	-0.80	-0.82	2.77
Unemployment mean	10.19	9.11	26.05	10.19	8.52	40.12	10.19	8.34	44.94	10.19	6.95	86.41	9.11	8.52	14.38
Commuters mean	-32.08	-46.66	17.85	-32.08	-40.64	10.30	-32.08	-29.15	-3.92	-32.08	-13.52	-27.89	-46.66	-40.64	-6.84
Age	55.38	55.46	-1.47	55.38	55.32	1.00	55.38	55.85	-8.00	55.38	55.57	-3.39	55.46	55.32	2.42
Number of children	0.24	0.22	3.40	0.24	0.25	-1.02	0.24	0.20	5.82	0.24	0.25	-1.42	0.22	0.25	-4.37
Male	0.49	0.49	0.00	0.49	0.51	-0.02	0.49	0.47	0.02	0.49	0.48	0.01	0.49	0.51	-0.02
Married	0.80	0.82	-0.02	0.80	0.81	-0.02	0.80	0.80	0.00	0.80	0.79	0.01	0.82	0.81	0.01
Divorced	0.09	0.07	0.02	0.09	0.09	0.00	0.09	0.08	0.01	0.09	0.10	0.00	0.07	0.09	-0.02
German	0.96	0.95	0.01	0.96	0.95	0.01	0.96	0.93	0.03	0.96	0.90	0.06	0.95	0.95	0.00
Other graduation diploma	0.06	0.05	0.01	0.06	0.06	0.00	0.06	0.08	-0.02	0.06	0.10	-0.04	0.05	0.06	-0.01
Intermediate secondary school	0.26	0.26	0.00	0.26	0.25	0.01	0.26	0.24	0.02	0.26	0.23	0.03	0.26	0.25	0.01
Upper secondary school	0.17	0.17	0.00	0.17	0.17	0.00	0.17	0.16	0.01	0.17	0.16	0.01	0.17	0.17	-0.01
No vocational degree	0.15	0.17	-0.01	0.15	0.13	0.03	0.15	0.15	0.00	0.15	0.19	-0.03	0.17	0.13	0.04
Degree below university	0.73	0.73	0.00	0.73	0.76	-0.03	0.73	0.75	-0.02	0.73	0.71	0.02	0.73	0.76	-0.03
University	0.22	0.21	0.01	0.22	0.20	0.02	0.22	0.19	0.03	0.22	0.19	0.03	0.21	0.20	0.01
Northern states	0.22	0.17	0.05	0.22	0.15	0.07	0.22	0.07	0.15	0.22	0.02	0.20	0.17	0.15	0.02
Western/Central states	0.07	0.08	-0.01	0.07	0.13	-0.06	0.07	0.16	-0.08	0.07	0.22	-0.14	0.08	0.13	-0.05
Southern states	0.11	0.17	-0.06	0.11	0.19	-0.08	0.11	0.30	-0.19	0.11	0.37	-0.26	0.17	0.19	-0.02
Eastern states	0.38	0.39	-0.01	0.38	0.30	0.09	0.38	0.29	0.10	0.38	0.13	0.25	0.39	0.30	0.10

	Block 2 vs 4			Block 2 vs 5			Block 3 vs 4			Block 3 vs 5			Block 4 vs 5		
	Mean 2	Mean 4	% bias	Mean 2	Mean 5	% bias	Mean 3	Mean 4	% bias	Mean 3	Mean 5	% bias	Mean 4	Mean 5	% bias
ln(Population per km2/100)	1.19	1.66	-40.23	1.19	1.88	-65.46	1.35	1.66	-26.27	1.35	1.88	-50.39	1.66	1.88	-22.00
ln(Average rain/100)	1.99	2.06	-31.58	1.99	2.12	-62.78	2.04	2.06	-10.34	2.04	2.12	-40.95	2.06	2.12	-28.51
ln(Proportion recreational area pc in sqm/100)	-0.80	-0.96	20.73	-0.80	-1.14	46.99	-0.82	-0.96	19.63	-0.82	-1.14	48.71	-0.96	-1.14	26.66
Unemployment mean	9.11	8.34	19.08	9.11	6.95	58.77	8.52	8.34	4.57	8.52	6.95	42.60	8.34	6.95	37.92
Commuters mean	-46.66	-29.15	-21.87	-46.66	-13.52	-45.71	-40.64	-29.15	-14.10	-40.64	-13.52	-36.61	-29.15	-13.52	-24.23
Age	55.46	55.85	-6.45	55.46	55.57	-1.86	55.32	55.85	-8.82	55.32	55.57	-4.33	55.85	55.57	4.74
Number of children	0.22	0.20	2.49	0.22	0.25	-4.93	0.25	0.20	6.73	0.25	0.25	-0.34	0.20	0.25	-7.42
Male	0.49	0.47	0.02	0.49	0.48	0.01	0.51	0.47	0.05	0.51	0.48	0.04	0.47	0.48	-0.01
Married	0.82	0.80	0.02	0.82	0.79	0.03	0.81	0.80	0.02	0.81	0.79	0.02	0.80	0.79	0.01
Divorced	0.07	0.08	-0.01	0.07	0.10	-0.03	0.09	0.08	0.01	0.09	0.10	-0.01	0.08	0.10	-0.02
German	0.95	0.93	0.02	0.95	0.90	0.05	0.95	0.93	0.02	0.95	0.90	0.05	0.93	0.90	0.03
Other graduation diploma	0.05	0.08	-0.03	0.05	0.10	-0.05	0.06	0.08	-0.01	0.06	0.10	-0.04	0.08	0.10	-0.02
Intermediate secondary school	0.26	0.24	0.02	0.26	0.23	0.02	0.25	0.24	0.01	0.25	0.23	0.01	0.24	0.23	0.01
Upper secondary school	0.17	0.16	0.01	0.17	0.16	0.01	0.17	0.16	0.01	0.17	0.16	0.02	0.16	0.16	0.00
No vocational degree	0.17	0.15	0.01	0.17	0.19	-0.02	0.13	0.15	-0.03	0.13	0.19	-0.06	0.15	0.19	-0.03
Degree below university	0.73	0.75	-0.02	0.73	0.71	0.02	0.76	0.75	0.02	0.76	0.71	0.06	0.75	0.71	0.04
University	0.21	0.19	0.02	0.21	0.19	0.02	0.20	0.19	0.01	0.20	0.19	0.01	0.19	0.19	0.00
Northern states	0.17	0.07	0.10	0.17	0.02	0.15	0.15	0.07	0.08	0.15	0.02	0.13	0.07	0.02	0.05
Western/Central states	0.08	0.16	-0.07	0.08	0.22	-0.13	0.13	0.16	-0.02	0.13	0.22	-0.08	0.16	0.22	-0.06
Southern states	0.17	0.30	-0.13	0.17	0.37	-0.20	0.19	0.30	-0.11	0.19	0.37	-0.18	0.30	0.37	-0.07
Eastern states	0.39	0.29	0.11	0.39	0.13	0.26	0.30	0.29	0.01	0.30	0.13	0.17	0.29	0.13	0.16

Sources: RDC of the Federal Statistical Office and Statistical Offices of the Länder (2018a; 2018b), SOEP as well as further data sources as discussed in Chapter 2, own calculations.

Table E.11: Balancing after step 2 of common support procedure - A3 sample

	Block 1 vs 2	Block 1 vs 3	Block 1 vs 4	Block 1 vs 5	Block 2 vs 3	Block 2 vs 4	Block 2 vs 5	Block 3 vs 4	Block 3 vs 5	Block 4 vs 5
ln(Population per km2/100)	11.08	10.10	3.47	0.52	-1.13	-8.48	-12.87	-7.39	-11.72	-3.67
ln(Average rain/100)	23.13	14.03	20.72	12.01	-9.73	-1.85	-12.04	7.58	-2.26	-9.80
ln(Proportion recreational area pc in sqm/100)	19.70	9.95	19.19	5.63	-10.87	-1.69	-16.03	9.77	-5.09	-15.32
Unemployment mean	16.55	17.57	5.06	11.42	1.10	-11.73	-6.96	-12.79	-8.14	6.00
Commuters mean	18.66	22.09	16.84	8.48	3.54	-3.32	-13.24	-7.08	-17.16	-10.76
Age	-2.53	-1.92	-7.36	-11.62	0.58	-4.77	-8.90	-5.31	-9.41	-4.00
Number of children	1.56	-2.03	4.05	-0.63	-3.63	2.58	-2.24	6.08	1.45	-4.77
Male	0.00	-0.03	0.01	0.00	-0.03	0.01	0.00	0.04	0.03	-0.01
Married	-0.01	0.00	0.01	0.00	0.01	0.02	0.01	0.01	0.00	-0.01
Divorced	0.01	-0.01	0.01	0.01	-0.02	-0.01	-0.01	0.01	0.01	0.00
German	0.01	0.00	0.00	0.00	-0.01	-0.01	-0.01	0.01	0.00	0.00
Other graduation diploma	0.01	0.00	0.00	0.00	-0.01	-0.01	-0.01	0.00	0.00	0.00
Intermediate secondary school	0.01	0.00	0.00	0.04	-0.02	-0.01	0.03	0.01	0.04	0.04
Upper secondary school	-0.01	-0.01	0.00	0.03	0.00	0.01	0.03	0.01	0.03	0.03
No vocational degree	-0.02	0.02	0.01	0.00	0.04	0.03	0.02	-0.01	-0.02	-0.01
Degree below university	0.02	-0.02	-0.02	-0.02	-0.04	-0.04	-0.04	0.00	0.00	0.00
University	-0.01	0.00	0.01	-0.03	0.01	0.01	-0.02	0.00	-0.03	-0.04
Northern states	0.02	0.01	0.04	0.02	-0.01	0.01	-0.01	0.02	0.01	-0.02
Western/Central states	-0.01	-0.03	-0.02	0.02	-0.02	-0.02	0.03	0.00	0.05	0.05
Southern states	-0.06	-0.03	-0.03	-0.01	0.03	0.03	0.05	0.00	0.02	0.02
Eastern states	-0.03	0.00	-0.05	-0.01	0.03	-0.02	0.02	-0.05	-0.01	0.04

Sources: RDC of the Federal Statistical Office and Statistical Offices of the Länder (2018a; 2018b), SOEP as well as further data sources as discussed in Chapter 2, own calculations.

Table E.12: Balancing after step 2 of common support procedure - A1 sample with lagged treatment and outcome

	Block 1 vs 2	Block 1 vs 3	Block 1 vs 4	Block 1 vs 5	Block 2 vs 3	Block 2 vs 4	Block 2 vs 5	Block 3 vs 4	Block 3 vs 5	Block 4 vs 5
ln(Population per km2/100)	30.38	16.64	60.25	-0.26	-14.78	31.09	-35.27	46.35	-19.55	-69.60
ln(Average rain/100)	56.09	33.73	44.53	43.43	-24.15	-9.63	-14.03	13.39	10.28	-3.70
ln(Proportion recreational area pc in sqm/100)	-2.33	3.13	-17.93	-11.65	5.28	-14.17	-8.18	-21.84	-15.56	7.35
Unemployment mean	-35.56	-22.79	-36.99	-51.94	14.79	0.73	-11.30	-15.01	-29.24	-13.02
Commuters mean	-3.85	-4.67	24.90	4.88	-0.89	27.78	8.74	27.87	9.45	-21.69
Dist.-weighted sports facility exp. 2002	-20.22	-25.18	-31.80	-26.37	-9.83	-20.85	-12.92	-12.29	-4.83	6.51
Age	-3.59	4.11	11.47	-12.13	7.65	15.01	-8.54	7.25	-16.06	-23.42
Number of children	48.27	31.66	42.42	61.54	-12.34	-3.63	15.28	8.51	25.42	17.84
Male	0.11	0.10	0.12	0.08	0.00	0.01	-0.02	0.02	-0.02	-0.04
Married	0.01	0.05	0.02	-0.02	0.04	0.01	-0.03	-0.04	-0.07	-0.04
Divorced	0.05	0.05	0.04	0.06	0.00	-0.01	0.01	-0.01	0.01	0.02
German	-0.16	-0.14	-0.19	-0.16	0.02	-0.03	0.00	-0.06	-0.02	0.03
Other graduation diploma	0.00	0.00	-0.01	-0.01	0.01	0.00	0.00	-0.01	-0.01	0.00
Intermediate secondary school	-0.02	0.00	0.01	0.06	0.01	0.02	0.08	0.01	0.06	0.05
Upper secondary school	-0.07	-0.09	-0.07	-0.20	-0.02	0.00	-0.13	0.02	-0.11	-0.13
Not yet graduated	0.02	-0.01	0.00	0.05	-0.03	-0.02	0.04	0.01	0.06	0.06
In education	-0.02	-0.05	0.00	-0.03	-0.03	0.02	-0.01	0.05	0.02	-0.02
No vocational degree	0.03	0.01	-0.01	0.14	-0.02	-0.04	0.11	-0.03	0.12	0.15
Degree below university	0.02	0.06	0.06	-0.04	0.04	0.04	-0.06	-0.01	-0.11	-0.10
University	-0.01	-0.02	0.01	-0.07	-0.01	0.01	-0.07	0.03	-0.05	-0.08
Northern states	-0.05	-0.13	-0.11	-0.03	-0.08	-0.06	0.01	0.01	0.09	0.08
Western/central states	0.12	0.11	0.09	0.13	-0.02	-0.03	0.01	-0.01	0.03	0.04
Southern states	0.04	0.11	0.14	0.18	0.07	0.10	0.14	0.03	0.07	0.04
Eastern states	-0.31	-0.17	-0.24	-0.24	0.14	0.07	0.07	-0.07	-0.07	0.00
Sports at least monthly	-0.15	-0.17	-0.12	-0.17	-0.02	0.03	-0.02	0.05	0.00	-0.05

Sources: RDC of the Federal Statistical Office and Statistical Offices of the Länder (2018a; 2018b), SOEP as well as further data sources as discussed in Chapter 2, own calculations.

Table E.13: Balancing after step 2 of common support procedure - A2 sample with lagged treatment and outcome

	Block 1 vs 2	Block 1 vs 3	Block 1 vs 4	Block 1 vs 5	Block 2 vs 3	Block 2 vs 4	Block 2 vs 5	Block 3 vs 4	Block 3 vs 5	Block 4 vs 5
ln(Population per km2/100)	33.91	15.78	21.89	6.01	-20.41	-12.60	-33.05	7.29	-11.90	-19.08
ln(Average rain/100)	32.15	34.20	12.84	17.96	1.01	-17.12	-14.87	-18.54	-16.41	3.64
ln(Proportion recreational area pc in sqm/100)	-4.98	-12.31	-6.33	-2.67	-6.28	-0.63	2.73	6.45	10.19	3.85
Unemployment mean	-11.45	-12.84	-15.65	-9.67	-1.18	-3.44	3.64	-2.25	5.12	8.14
Commuters mean	-2.38	2.75	-4.07	-33.01	5.00	-1.26	-28.87	-7.05	-35.42	-33.95
Dist.-weighted sports facility exp. 2002	-16.64	-23.68	-19.82	-4.43	-12.74	-7.03	14.54	4.17	23.08	18.25
Age	-13.42	-10.59	-4.95	13.39	2.65	8.90	27.27	6.04	24.16	18.94
Number of children	-12.94	-10.63	-14.39	10.05	1.99	-1.89	23.42	-3.77	20.77	24.57
Male	0.00	-0.07	-0.02	-0.04	-0.07	-0.02	-0.04	0.05	0.03	-0.02
Married	-0.01	0.01	0.02	0.07	0.02	0.03	0.08	0.01	0.06	0.05
Divorced	0.06	0.06	0.03	0.01	0.00	-0.03	-0.05	-0.04	-0.05	-0.02
German	-0.02	-0.02	-0.03	-0.05	0.00	-0.01	-0.03	-0.01	-0.03	-0.02
Other graduation diploma	-0.01	-0.01	-0.03	0.00	0.01	-0.02	0.01	-0.02	0.01	0.03
Intermediate secondary school	-0.12	-0.08	-0.09	-0.10	0.04	0.04	0.03	-0.01	-0.02	-0.01
Upper secondary school	0.11	0.10	0.15	0.06	-0.01	0.04	-0.05	0.05	-0.04	-0.09
No vocational degree	0.05	0.03	0.04	0.04	-0.02	-0.02	-0.01	0.00	0.01	0.01
Degree below university	0.01	0.00	-0.05	0.01	-0.01	-0.06	0.01	-0.06	0.01	0.07
University	-0.09	-0.02	0.02	-0.13	0.07	0.10	-0.04	0.04	-0.10	-0.14
Northern states	-0.05	-0.15	-0.14	0.01	-0.10	-0.09	0.06	0.01	0.16	0.15
Western/Central states	0.05	0.02	0.01	0.03	-0.04	-0.04	-0.02	-0.01	0.01	0.02
Southern states	-0.04	0.08	0.08	0.01	0.12	0.12	0.05	-0.01	-0.07	-0.07
Eastern states	-0.13	-0.08	-0.10	-0.06	0.05	0.03	0.07	-0.02	0.02	0.04
Sports at least monthly	-0.01	0.00	0.04	-0.04	0.00	0.04	-0.03	0.04	-0.04	-0.08

Sources: RDC of the Federal Statistical Office and Statistical Offices of the Länder (2018a; 2018b), SOEP as well as further data sources as discussed in Chapter 2, own calculations.

Table E.14: Balancing after step 2 of common support procedure - A3 sample with lagged treatment and outcome

	Block 1 vs 2	Block 1 vs 3	Block 1 vs 4	Block 1 vs 5	Block 2 vs 3	Block 2 vs 4	Block 2 vs 5	Block 3 vs 4	Block 3 vs 5	Block 4 vs 5
ln(Population per km2/100)	47.68	54.62	44.19	24.91	7.07	-5.03	-30.64	-12.35	-38.98	-26.04
ln(Average rain/100)	52.46	20.77	25.04	15.33	-33.35	-26.54	-39.45	5.38	-5.88	-11.03
ln(Proportion recreational area pc in sqm/100)	-5.51	-23.37	-8.54	12.73	-16.45	-2.54	17.95	14.82	38.96	22.26
Unemployment mean	7.08	16.59	6.82	5.04	9.68	-0.29	-2.82	-10.01	-13.50	-2.51
Commuters mean	9.46	17.19	12.11	-11.82	7.46	1.65	-21.51	-6.43	-29.91	-26.22
Dist.-weighted sports facility exp. 2002	-9.83	-7.21	-10.37	-2.28	3.34	-2.33	6.78	-4.88	4.31	7.70
Age	-33.86	-29.92	-26.30	-36.38	3.60	7.40	-1.68	3.74	-5.35	-9.24
Number of children	67.56	67.18	74.81	83.39	4.24	6.55	18.48	1.77	13.09	12.41
Male	0.14	0.10	0.15	0.19	-0.04	0.01	0.05	0.05	0.09	0.04
Married	0.01	0.02	0.07	-0.01	0.01	0.06	-0.02	0.05	-0.02	-0.07
Divorced	0.01	0.01	0.02	-0.03	-0.01	0.01	-0.05	0.01	-0.04	-0.05
German	0.00	-0.02	-0.02	0.00	-0.02	-0.03	-0.01	-0.01	0.01	0.02
Other graduation diploma	-0.02	-0.02	-0.07	-0.02	0.00	-0.06	0.00	-0.06	0.00	0.06
Intermediate secondary school	-0.05	-0.08	0.00	0.02	-0.03	0.05	0.07	0.08	0.10	0.02
Upper secondary school	0.04	0.08	0.12	0.09	0.04	0.08	0.05	0.05	0.02	-0.03
No vocational degree	-0.02	0.03	-0.03	-0.04	0.05	-0.01	-0.03	-0.06	-0.07	-0.02
Degree below university	0.07	0.02	0.02	0.05	-0.05	-0.05	-0.01	0.00	0.04	0.04
University	-0.10	-0.06	0.02	-0.02	0.03	0.12	0.08	0.08	0.05	-0.04
Northern states	-0.05	-0.09	-0.04	-0.01	-0.04	0.00	0.04	0.04	0.08	0.03
Western / central states	-0.05	-0.06	-0.22	-0.04	-0.01	-0.17	0.01	-0.16	0.02	0.18
Southern states	-0.05	-0.07	-0.04	0.01	-0.02	0.01	0.06	0.03	0.08	0.05
Eastern states	-0.13	-0.04	-0.01	-0.08	0.09	0.12	0.04	0.03	-0.05	-0.08
Sports at least monthly	0.16	0.15	0.17	0.13	-0.01	0.02	-0.03	0.03	-0.02	-0.04

Sources: RDC of the Federal Statistical Office and Statistical Offices of the Länder (2018a; 2018b), SOEP as well as further data sources as discussed in Chapter 2, own calculations.

Table E.15: Balancing before GPS adjustment - A2 sample - small municipalities

	Block 1 vs 2			Block 1 vs 3			Block 1 vs 4			Block 1 vs 5			Block 2 vs 3		
	Mean 1	Mean 2	% bias	Mean 1	Mean 3	% bias	Mean 1	Mean 4	% bias	Mean 1	Mean 5	% bias	Mean 2	Mean 3	% bias
ln(Population per km2/100)	0.15	0.28	-16.13	0.15	0.55	-48.31	0.15	0.56	-48.16	0.15	0.96	-100.24	0.28	0.55	-34.12
ln(Average rain/100)	2.03	2.03	0.94	2.03	2.01	10.57	2.03	2.12	-38.48	2.03	2.14	-56.29	2.03	2.01	9.47
ln(Proportion recreational area p.c. in sqm/100)	-0.52	-0.53	2.21	-0.52	-0.57	6.96	-0.52	-0.77	30.94	-0.52	-1.02	61.21	-0.53	-0.57	4.51
Unemployment mean	8.21	8.57	-7.76	8.21	8.06	3.22	8.21	7.56	14.69	8.21	5.88	58.13	8.57	8.06	11.27
Commuters mean	-96.19	-102.83	6.42	-96.19	-90.49	-5.66	-96.19	-60.83	-39.46	-96.19	-57.18	-47.20	-102.83	-90.49	-11.34
Age	37.97	38.36	-9.06	37.97	38.80	-19.50	37.97	37.77	4.95	37.97	37.75	5.08	38.36	38.80	-10.59
Number of children	1.31	1.41	-9.27	1.31	1.33	-1.98	1.31	1.53	-19.34	1.31	1.39	-6.90	1.41	1.33	7.84
Male	0.48	0.48	0.00	0.48	0.47	0.02	0.48	0.47	0.01	0.48	0.47	0.01	0.48	0.47	0.01
Married	0.75	0.80	-0.06	0.75	0.78	-0.03	0.75	0.76	-0.02	0.75	0.79	-0.04	0.80	0.78	0.03
Divorced	0.08	0.06	0.02	0.08	0.08	0.00	0.08	0.07	0.00	0.08	0.08	0.00	0.06	0.08	-0.02
German nationality	0.99	0.98	0.01	0.99	0.98	0.01	0.99	0.96	0.02	0.99	0.88	0.11	0.98	0.98	0.00
Other graduation diploma	0.03	0.03	-0.01	0.03	0.04	-0.01	0.03	0.04	-0.02	0.03	0.07	-0.05	0.03	0.04	-0.01
Intermediate secondary school	0.45	0.56	-0.11	0.45	0.48	-0.03	0.45	0.47	-0.02	0.45	0.40	0.05	0.56	0.48	0.08
Upper secondary school	0.16	0.17	-0.01	0.16	0.20	-0.04	0.16	0.19	-0.03	0.16	0.22	-0.06	0.17	0.20	-0.02
No vocational degree	0.11	0.08	0.03	0.11	0.11	0.00	0.11	0.10	0.01	0.11	0.16	-0.05	0.08	0.11	-0.03
Degree below university	0.82	0.84	-0.01	0.82	0.79	0.03	0.82	0.81	0.01	0.82	0.76	0.06	0.84	0.79	0.04
University degree	0.14	0.20	-0.06	0.14	0.17	-0.03	0.14	0.17	-0.03	0.14	0.15	-0.01	0.20	0.17	0.03
Northern states	0.24	0.18	0.05	0.24	0.09	0.14	0.24	0.13	0.10	0.24	0.04	0.20	0.18	0.09	0.09
Western/Central states	0.16	0.12	0.04	0.16	0.21	-0.05	0.16	0.15	0.01	0.16	0.25	-0.09	0.12	0.21	-0.09
Southern states	0.28	0.23	0.05	0.28	0.28	0.00	0.28	0.31	-0.03	0.28	0.50	-0.22	0.23	0.28	-0.05
Eastern states	0.29	0.38	-0.09	0.29	0.36	-0.07	0.29	0.33	-0.04	0.29	0.17	0.12	0.38	0.36	0.02

	Block 2 vs 4			Block 2 vs 5			Block 3 vs 4			Block 3 vs 5			Block 4 vs 5		
	Mean 2	Mean 4	% bias	Mean 2	Mean 5	% bias	Mean 3	Mean 4	% bias	Mean 3	Mean 5	% bias	Mean 4	Mean 5	% bias
ln(Population per km2/100)	0.28	0.56	-34.38	0.28	0.96	-88.01	0.55	0.56	-1.38	0.55	0.96	-51.02	0.56	0.96	-47.91
ln(Average rain/100)	2.03	2.12	-38.84	2.03	2.14	-56.37	2.01	2.12	-47.61	2.01	2.14	-66.53	2.12	2.14	-11.36
ln(Proportion recreational area p.c. in sqm/100)	-0.53	-0.77	27.76	-0.53	-1.02	56.98	-0.57	-0.77	24.13	-0.57	-1.02	54.28	-0.77	-1.02	29.13
Unemployment mean	8.57	7.56	23.04	8.57	5.88	67.47	8.06	7.56	11.86	8.06	5.88	57.31	7.56	5.88	46.14
Commuters mean	-102.83	-60.83	-42.59	-102.83	-57.18	-49.44	-90.49	-60.83	-30.96	-90.49	-57.18	-37.28	-60.83	-57.18	-4.76
Age	38.36	37.77	14.56	38.36	37.75	14.31	38.80	37.77	25.55	38.80	37.75	24.88	37.77	37.75	0.33
Number of children	1.41	1.53	-10.88	1.41	1.39	1.97	1.33	1.53	-18.61	1.33	1.39	-5.38	1.53	1.39	12.21
Male	0.48	0.47	0.01	0.48	0.47	0.01	0.47	0.47	0.00	0.47	0.47	0.00	0.47	0.47	0.00
Married	0.80	0.76	0.04	0.80	0.79	0.01	0.78	0.76	0.02	0.78	0.79	-0.01	0.76	0.79	-0.03
Divorced	0.06	0.07	-0.01	0.06	0.08	-0.02	0.08	0.07	0.01	0.08	0.08	0.00	0.07	0.08	-0.01
German nationality	0.98	0.96	0.01	0.98	0.88	0.10	0.98	0.96	0.01	0.98	0.88	0.10	0.96	0.88	0.08
Other graduation diploma	0.03	0.04	-0.01	0.03	0.07	-0.04	0.04	0.04	0.00	0.04	0.07	-0.03	0.04	0.07	-0.03
Intermediate secondary school	0.56	0.47	0.09	0.56	0.40	0.16	0.48	0.47	0.01	0.48	0.40	0.08	0.47	0.40	0.07
Upper secondary school	0.17	0.19	-0.02	0.17	0.22	-0.05	0.20	0.19	0.01	0.20	0.22	-0.02	0.19	0.22	-0.03
No vocational degree	0.08	0.10	-0.02	0.08	0.16	-0.08	0.11	0.10	0.01	0.11	0.16	-0.05	0.10	0.16	-0.06
Degree below university	0.84	0.81	0.03	0.84	0.76	0.07	0.79	0.81	-0.02	0.79	0.76	0.03	0.81	0.76	0.05
University degree	0.20	0.17	0.03	0.20	0.15	0.06	0.17	0.17	0.00	0.17	0.15	0.02	0.17	0.15	0.02
Northern states	0.18	0.13	0.05	0.18	0.04	0.15	0.09	0.13	-0.04	0.09	0.04	0.06	0.13	0.04	0.10
Western/Central states	0.12	0.15	-0.03	0.12	0.25	-0.13	0.21	0.15	0.06	0.21	0.25	-0.03	0.15	0.25	-0.10
Southern states	0.23	0.31	-0.09	0.23	0.50	-0.28	0.28	0.31	-0.04	0.28	0.50	-0.22	0.31	0.50	-0.19
Eastern states	0.38	0.33	0.05	0.38	0.17	0.21	0.36	0.33	0.03	0.36	0.17	0.18	0.33	0.17	0.15

Sources: RDC of the Federal Statistical Office and Statistical Offices of the Länder (2018a; 2018b), SOEP as well as further data sources as discussed in Chapter 2, own calculations.

Table E.16: Balancing after step 2 of common support procedure - A2 sample - small municipalities

	Block 1 vs 2	Block 1 vs 3	Block 1 vs 4	Block 1 vs 5	Block 2 vs 3	Block 2 vs 4	Block 2 vs 5	Block 3 vs 4	Block 3 vs 5	Block 4 vs 5
ln(Population per km2/100)	-7.32	-17.53	2.01	-9.41	-10.90	9.20	-2.26	19.08	8.61	-11.21
ln(Average rain/100)	8.85	30.36	-3.02	20.13	21.05	-10.69	11.19	-29.36	-9.51	20.53
ln(Proportion recreational area pc in sqm/100)	6.06	-1.03	9.83	-18.44	-7.03	3.64	-23.70	10.80	-17.39	-27.39
Unemployment mean	-9.93	-18.20	-18.65	-17.28	-7.96	-8.11	-5.76	0.10	3.36	3.39
Commuters mean	4.67	5.84	-24.29	-5.61	0.97	-26.97	-10.25	-28.86	-11.78	22.36
Age	-23.03	-26.39	-5.21	-19.81	-3.31	18.87	3.10	22.38	6.40	-15.51
Number of children	-9.85	-1.46	-13.09	15.99	9.01	-3.74	26.49	-12.48	18.45	29.17
Male	0.01	0.00	-0.02	0.02	-0.01	-0.03	0.01	-0.01	0.02	0.04
Married	-0.04	0.00	0.04	0.01	0.03	0.07	0.05	0.04	0.02	-0.02
Divorced	0.01	-0.01	-0.01	-0.03	-0.02	-0.02	-0.04	-0.01	-0.03	-0.02
German nationality	0.02	0.00	-0.01	-0.01	-0.02	-0.02	-0.02	0.00	0.00	0.00
Other graduation diploma	-0.01	0.00	0.01	0.02	0.00	0.01	0.02	0.01	0.02	0.01
Intermediate secondary school	-0.10	-0.03	-0.04	-0.07	0.08	0.06	0.03	-0.02	-0.04	-0.03
Upper secondary school	-0.03	-0.04	-0.01	-0.03	0.00	0.03	0.01	0.03	0.01	-0.02
No vocational degree	0.02	-0.01	0.00	0.03	-0.03	-0.02	0.01	0.01	0.04	0.03
Degree below university	0.01	0.06	0.03	-0.02	0.05	0.03	-0.02	-0.02	-0.07	-0.05
University	-0.06	-0.07	-0.04	-0.06	0.00	0.03	0.01	0.03	0.01	-0.02
Northern states	0.02	0.08	0.02	0.03	0.05	0.00	0.01	-0.06	-0.04	0.01
Western/Central states	0.08	0.02	0.08	0.01	-0.06	0.00	-0.07	0.06	-0.01	-0.07
Southern states	0.03	0.04	0.07	0.09	0.01	0.04	0.06	0.02	0.05	0.02
Eastern states	-0.08	-0.12	-0.12	-0.12	-0.04	-0.04	-0.03	0.00	0.01	0.01

Sources: RDC of the Federal Statistical Office and Statistical Offices of the Länder (2018a; 2018b), SOEP as well as further data sources as discussed in Chapter 2, own calculations.

Table E.17: Balancing before GPS adjustment - A2 sample - large municipalities

	Block 1 vs 2			Block 1 vs 3			Block 1 vs 4			Block 1 vs 5			Block 2 vs 3		
	Mean 1	Mean 2	% bias	Mean 1	Mean 3	% bias	Mean 1	Mean 4	% bias	Mean 1	Mean 5	% bias	Mean 2	Mean 3	% bias
ln(Population per km2/100)	2.21	2.17	5.67	2.21	2.15	7.38	2.21	2.50	-35.14	2.21	2.34	-16.98	2.17	2.15	1.85
ln(Average rain/100)	2.03	2.04	-3.42	2.03	2.08	-26.19	2.03	2.03	-1.01	2.03	2.14	-53.49	2.04	2.08	-23.16
ln(Proportion recreational area pc in sqm/100)	-0.88	-1.06	31.77	-0.88	-1.05	29.48	-0.88	-1.16	53.05	-0.88	-1.24	64.98	-1.06	-1.05	-1.54
Unemployment mean	10.17	9.54	18.49	10.17	8.63	42.59	10.17	7.84	69.32	10.17	7.00	99.02	9.54	8.63	26.08
Commuters mean	11.84	9.03	9.49	11.84	-0.15	36.48	11.84	7.66	11.86	11.84	10.97	2.51	9.03	-0.15	26.84
Age	38.08	38.06	0.31	38.08	37.85	5.19	38.08	37.74	7.94	38.08	37.78	7.05	38.06	37.85	4.98
Number of children	1.21	1.28	-6.94	1.21	1.20	0.96	1.21	1.06	13.64	1.21	1.15	5.42	1.28	1.20	7.90
Male	0.46	0.46	0.00	0.46	0.50	-0.04	0.46	0.49	-0.03	0.46	0.47	-0.01	0.46	0.50	-0.04
Married	0.68	0.69	-0.02	0.68	0.74	-0.06	0.68	0.66	0.01	0.68	0.68	-0.01	0.69	0.74	-0.04
Divorced	0.10	0.13	-0.03	0.10	0.07	0.03	0.10	0.08	0.02	0.10	0.09	0.01	0.13	0.07	0.06
German nationality	0.91	0.92	0.00	0.91	0.90	0.01	0.91	0.85	0.06	0.91	0.88	0.03	0.92	0.90	0.01
Other graduation diploma	0.05	0.08	-0.02	0.05	0.07	-0.02	0.05	0.08	-0.02	0.05	0.09	-0.04	0.08	0.07	0.01
Intermediate secondary school	0.42	0.41	0.01	0.42	0.37	0.05	0.42	0.35	0.07	0.42	0.34	0.08	0.41	0.37	0.04
Upper secondary school	0.26	0.22	0.04	0.26	0.31	-0.05	0.26	0.26	0.00	0.26	0.34	-0.08	0.22	0.31	-0.09
No vocational degree	0.12	0.10	0.02	0.12	0.11	0.02	0.12	0.17	-0.04	0.12	0.13	-0.01	0.10	0.11	-0.01
Degree below university	0.74	0.77	-0.03	0.74	0.75	-0.01	0.74	0.69	0.05	0.74	0.71	0.03	0.77	0.75	0.02
University	0.23	0.22	0.02	0.23	0.21	0.02	0.23	0.22	0.02	0.23	0.27	-0.03	0.22	0.21	0.01
Northern states	0.18	0.21	-0.03	0.18	0.13	0.05	0.18	0.03	0.15	0.18	0.01	0.17	0.21	0.13	0.08
Western/central states	0.05	0.04	0.02	0.05	0.10	-0.05	0.05	0.26	-0.20	0.05	0.19	-0.13	0.04	0.10	-0.06
Southern states	0.11	0.15	-0.04	0.11	0.09	0.02	0.11	0.34	-0.23	0.11	0.39	-0.29	0.15	0.09	0.05
Eastern states	0.30	0.27	0.03	0.30	0.19	0.11	0.30	0.11	0.18	0.30	0.09	0.21	0.27	0.19	0.08

	Block 2 vs 4			Block 2 vs 5			Block 3 vs 4			Block 3 vs 5			Block 4 vs 5		
	Mean 2	Mean 4	% bias	Mean 2	Mean 5	% bias	Mean 3	Mean 4	% bias	Mean 3	Mean 5	% bias	Mean 4	Mean 5	% bias
ln(Population per km2/100)	2.17	2.50	-43.39	2.17	2.34	-24.76	2.15	2.50	-45.00	2.15	2.34	-26.58	2.50	2.34	21.82
ln(Average rain/100)	2.04	2.03	2.39	2.04	2.14	-51.17	2.08	2.03	25.15	2.08	2.14	-30.18	2.03	2.14	-52.52
ln(Proportion recreational area pc in sqm/100)	-1.06	-1.16	19.97	-1.06	-1.24	35.12	-1.05	-1.16	20.95	-1.05	-1.24	35.61	-1.16	-1.24	18.56
Unemployment mean	9.54	7.84	52.72	9.54	7.00	83.01	8.63	7.84	22.95	8.63	7.00	49.44	7.84	7.00	27.68
Commuters mean	9.03	7.66	3.74	9.03	10.97	-5.40	-0.15	7.66	-19.99	-0.15	10.97	-28.74	7.66	10.97	-8.12
Age	38.06	37.74	7.78	38.06	37.78	6.87	37.85	37.74	2.46	37.85	37.78	1.69	37.74	37.78	-0.75
Number of children	1.28	1.06	20.71	1.28	1.15	12.42	1.20	1.06	12.66	1.20	1.15	4.45	1.06	1.15	-8.24
Male	0.46	0.49	-0.03	0.46	0.47	-0.01	0.50	0.49	0.01	0.50	0.47	0.03	0.49	0.47	0.02
Married	0.69	0.66	0.03	0.69	0.68	0.01	0.74	0.66	0.07	0.74	0.68	0.05	0.66	0.68	-0.02
Divorced	0.13	0.08	0.05	0.13	0.09	0.04	0.07	0.08	-0.01	0.07	0.09	-0.02	0.08	0.09	-0.01
German nationality	0.92	0.85	0.06	0.92	0.88	0.04	0.90	0.85	0.05	0.90	0.88	0.02	0.85	0.88	-0.03
Other graduation diploma	0.08	0.08	0.00	0.08	0.09	-0.01	0.07	0.08	-0.01	0.07	0.09	-0.02	0.08	0.09	-0.01
Intermediate secondary school	0.41	0.35	0.06	0.41	0.34	0.07	0.37	0.35	0.02	0.37	0.34	0.03	0.35	0.34	0.01
Upper secondary school	0.22	0.26	-0.05	0.22	0.34	-0.12	0.31	0.26	0.04	0.31	0.34	-0.03	0.26	0.34	-0.08
No vocational degree	0.10	0.17	-0.07	0.10	0.13	-0.03	0.11	0.17	-0.06	0.11	0.13	-0.02	0.17	0.13	0.04
Degree below university	0.77	0.69	0.08	0.77	0.71	0.07	0.75	0.69	0.07	0.75	0.71	0.05	0.69	0.71	-0.02
University	0.22	0.22	0.00	0.22	0.27	-0.05	0.21	0.22	-0.01	0.21	0.27	-0.06	0.22	0.27	-0.05
Northern states	0.21	0.03	0.18	0.21	0.01	0.20	0.13	0.03	0.09	0.13	0.01	0.12	0.03	0.01	0.03
Western/central states	0.04	0.26	-0.22	0.04	0.19	-0.15	0.10	0.26	-0.15	0.10	0.19	-0.08	0.26	0.19	0.07
Southern states	0.15	0.34	-0.19	0.15	0.39	-0.25	0.09	0.34	-0.24	0.09	0.39	-0.30	0.34	0.39	-0.06
Eastern states	0.27	0.11	0.15	0.27	0.09	0.18	0.19	0.11	0.07	0.19	0.09	0.10	0.11	0.09	0.02

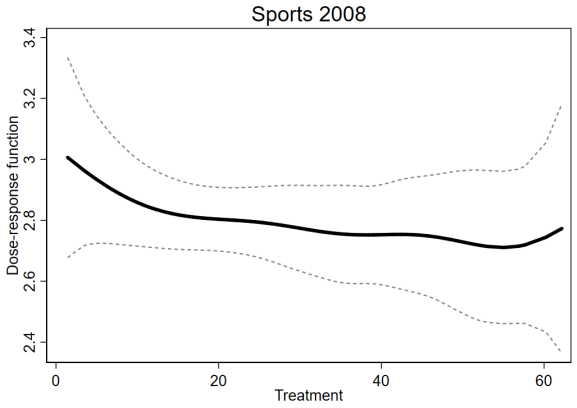
Sources: RDC of the Federal Statistical Office and Statistical Offices of the Länder (2018a; 2018b), SOEP as well as further data sources as discussed in Chapter 2, own calculations.

Table E.18: Balancing after step 2 of common support procedure - A2 sample - large municipalities

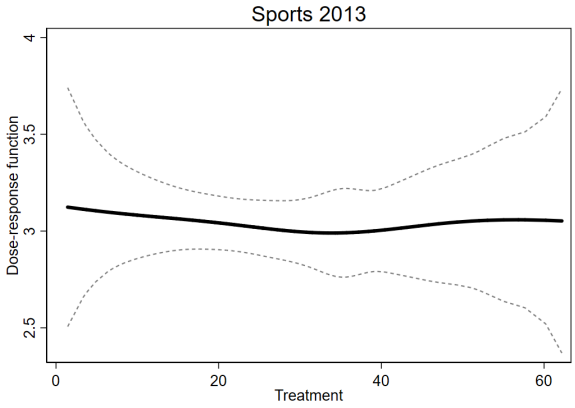
	Block 1 vs 2	Block 1 vs 3	Block 1 vs 4	Block 1 vs 5	Block 2 vs 3	Block 2 vs 4	Block 2 vs 5	Block 3 vs 4	Block 3 vs 5	Block 4 vs 5
ln(Population per km2/100)	15.91	8.67	-24.01	3.32	-7.68	-42.17	-14.46	-34.52	-6.25	30.50
ln(Average rain/100)	-10.32	-3.13	34.85	7.79	7.85	45.93	18.65	40.34	11.60	-28.22
ln(Proportion recreational area pc in sqm/100)	15.27	1.01	2.99	-22.23	-14.75	-14.76	-40.33	1.94	-24.15	-29.62
Unemployment mean	12.40	14.36	-4.43	-13.56	2.80	-17.66	-27.98	-19.34	-28.94	-9.45
Commuters mean	6.77	35.12	16.81	2.43	27.90	10.73	-3.23	-14.39	-27.67	-12.47
Age	6.61	8.53	6.83	5.60	2.22	0.20	-0.92	-2.04	-3.07	-1.12
Number of children	5.97	3.01	13.15	3.86	-2.97	6.97	-2.20	10.06	0.81	-9.35
Male	-0.01	-0.04	-0.02	-0.04	-0.03	-0.01	-0.03	0.02	0.00	-0.02
Married	0.02	-0.04	0.01	0.02	-0.06	-0.01	0.01	0.06	0.07	0.01
Divorced	-0.06	0.01	0.01	-0.01	0.07	0.07	0.04	0.00	-0.02	-0.02
German nationality	-0.01	-0.01	0.03	-0.03	0.00	0.03	-0.02	0.04	-0.02	-0.06
Other graduation diploma	-0.05	-0.03	-0.02	0.00	0.01	0.03	0.04	0.01	0.03	0.02
Intermediate secondary school	-0.05	0.03	0.00	-0.10	0.07	0.05	-0.05	-0.03	-0.13	-0.10
Upper secondary school	0.12	0.00	0.09	0.06	-0.11	-0.02	-0.05	0.09	0.06	-0.03
No vocational degree	0.00	0.01	-0.06	0.02	0.02	-0.06	0.02	-0.08	0.01	0.09
Degree below university	-0.03	-0.03	0.04	-0.07	-0.01	0.06	-0.04	0.07	-0.03	-0.10
University	0.07	0.08	0.05	0.06	0.01	-0.03	-0.02	-0.03	-0.02	0.01
Northern states	-0.02	-0.02	0.04	0.07	0.00	0.06	0.09	0.06	0.09	0.02
Western/Central states	0.03	-0.05	-0.08	0.02	-0.08	-0.11	-0.01	-0.02	0.07	0.10
Southern states	-0.02	0.13	0.02	0.02	0.14	0.03	0.04	-0.11	-0.10	0.01
Eastern states	0.01	0.03	-0.02	-0.13	0.02	-0.03	-0.14	-0.06	-0.16	-0.11

Sources: RDC of the Federal Statistical Office and Statistical Offices of the Länder (2018a; 2018b), SOEP as well as further data sources as discussed in Chapter 2, own calculations.

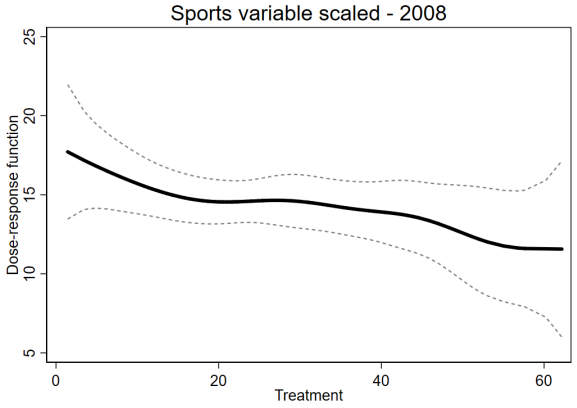
Figure E.8: Additional results A1 sample (17-30 years of age)



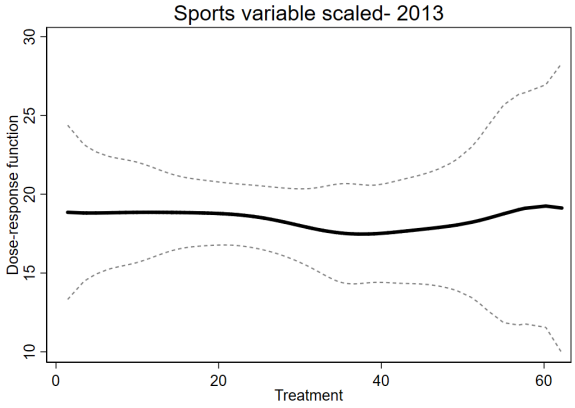
(a) DRF sports 2008



(b) DRF sports 2013



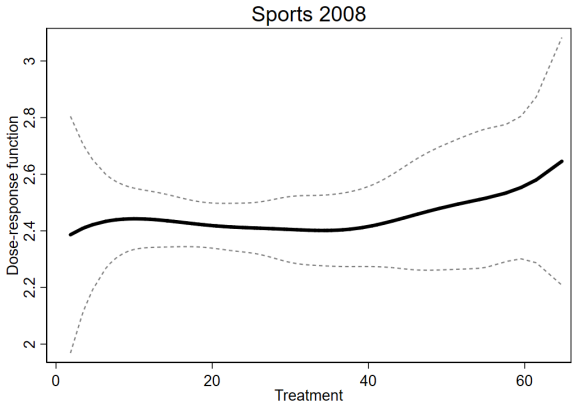
(c) DRF sports variable scaled - 2008



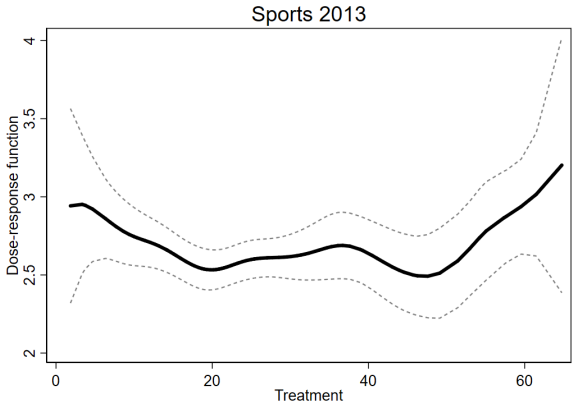
(d) DRF sports variable scaled - 2013

Sources: RDC of the Federal Statistical Office and Statistical Offices of the Länder (2018a; 2018b), SOEP as well as further data sources as discussed in Chapter 2, own calculations.

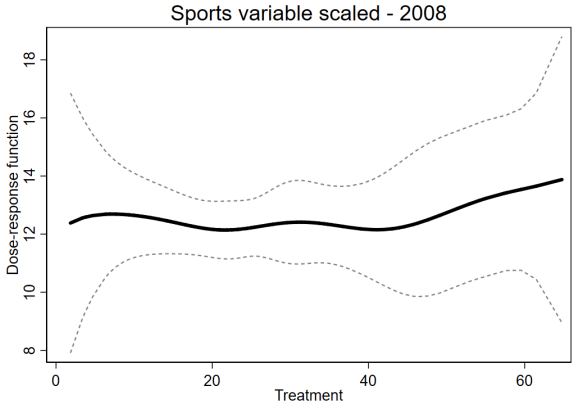
Figure E.9: Additional results A3 sample (46-65 years of age)



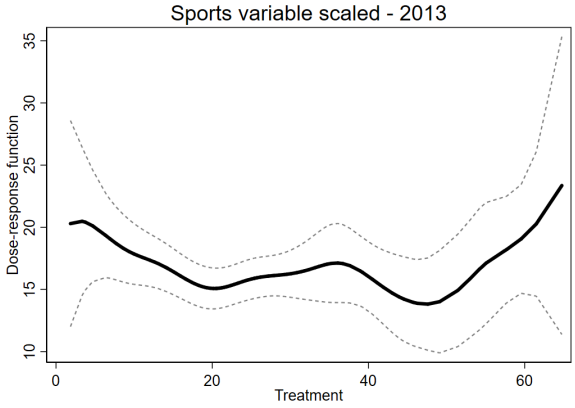
(a) DRF sports 2008



(b) DRF sports 2013



(c) DRF sports variable scaled - 2008



(d) DRF sports variable scaled - 2013

Sources: RDC of the Federal Statistical Office and Statistical Offices of the Länder (2018a; 2018b), SOEP as well as further data sources as discussed in Chapter 2, own calculations.