Essays in Empirical Economics

DISSERTATION
of the University of St.Gallen,
School of Management,
Economics, Law, Social Sciences
and International Affairs
to obtain the title of
Doctor of Philosophy in Economics and Finance

submitted by
Carina Steckenleiter
from Germany

Approved on the application of
Prof. Dr. Michael Lechner
and
Prof. Paul Downward, PhD

Dissertation no. 4969
Difo-Druck GmbH, Untersiemau 2020

Essays in Empirical Economics

DISSERTATION
of the University of St.Gallen,
School of Management,
Economics, Law, Social Sciences
and International Affairs
to obtain the title of
Doctor of Philosophy in Economics and Finance

submitted by
Carina Steckenleiter
from Germany

Approved on the application of
Prof. Dr. Michael Lechner
and
Prof. Paul Downward, PhD
Prof. Francesco Audrino, PhD

Dissertation no. 4969
Difo-Druck GmbH, Untersiemau 2020

The University of St.Gallen, School of Management, Economics, Law, Social Sciences and International Affairs hereby consents to the printing of the present dissertation, without hereby expressing any opinion on the views herein expressed.

St.Gallen, November 15, 2019

The President:

Prof. Dr. Thomas Bieger

Acknowledgments

First and foremost, I want to thank Michael Lechner for his excellent PhD-supervision and the possibility to write this dissertation under his supervision. I am very grateful for his constant advice, fast feedback, the many things I was able to learn during my PhD-studies, and the possibility to work on a diverse set of topics and projects.

Also, I want to thank my colleagues at the Swiss Institute for Empirical Economic Research and at the Center for Disability and Integration. It has been a great privilege and inspiration to be surrounded by all these nice and highly motivated people. I particularly want to thank my long-time office mates Michael Knaus and Michael Zimmert for discussing research and other things, providing many moments of laughter, and creating a positive office environment. I also would like to thank Michael Knaus and Michael Zimmert for proof-reading final versions of this thesis. A big thanks also belongs to Conny David for her great help and support in all sorts of administrative and organizational matters.

I am very grateful to my co-authors Iris Grant, Iris Kesternich, Michael Lechner, Tim Pawlowski, Bettina Siflinger, Ute Schüttoff, James P. Smith, Tim Wallrafen, and Joachim Winter as well everything I was able to learn from working with them. A special thanks goes to Iris Kesternich, Bettina Siflinger, and Joachim Winter. Working for Iris Kesternich as a student assistant in Munich and writing my master thesis at the chair of Joachim Winter inspired me to write a dissertation in the first place. I also want to thank the three of them for their valuable comments on my single authored paper and for their advice and support throughout my dissertation.

Parts of this dissertation are based on analyses conducted during research visits at the Research Data Center of the SOEP in Berlin and the German Microdata Lab in Mannheim. Dr. Jan Goebel from SOEP-DIW provided valuable insights on SOEP regional data and support with IT-Infrastructure during my research visits in Berlin. Michaela Engelmann brightened up every day with her positive attitude and with cake, cookies, and coffee. Kim and Christian took great care of me throughout my research visits in Berlin. Dr. Bernhard Schimpl-Neimanns from the German Microdata Lab took a lot of time answering my questions regarding Census and Microcensus data and provided valuable insights and background information on these data sets. I would like to thank all of them for their great help and support. I also thank the Swiss National Science Foundation for funding through grant SNF 100018E-161598.

I am indebted to the employees at various statistical offices, archives, and libraries across Germany who provided assistance in the collection of various data for chapters of

this dissertation. Their readiness to help is highly appreciated.

I thank Ferdinand Kramer for guidance and valuable advice throughout all my studies. I am very grateful to my family and especially my parents for their unconditional support in all conceivable ways. Finally, I want to thank Nico for his tremendously positive attitude and his invaluable support throughout my PhD-studies.

St.Gallen, December 2019 Carina Steckenleiter

Contents

List of Tables **v**

List of Figures **ix**

1 Effects of vocational training for women in a society with traditional role models – Evidence from West Germany **1**

1.1 Introduction . . . 2
1.2 Related literature . . . 5
1.3 Vocational training requirement in Germany . . . 6
1.4 Identification and estimation strategy . . . 10
1.5 Data . . . 11
1.5.1 Census 1970 . . . 11
1.5.2 Microcensus Supplementary Survey 1971 (MZU 71) . . . 12
1.5.3 Summary statistics . . . 14
1.6 Results . . . 15
1.6.1 Social status of husband - MZU 71 . . . 16
1.6.2 Labor market outcomes - Census 1970 . . . 19
1.7 Robustness and channels . . . 21
1.7.1 Robustness . . . 21
1.7.2 Channels . . . 25
1.8 Conclusions . . . 27

2 Unbalanced sex ratios in Germany caused by World War II and their effect on fertility: A life cycle perspective **29**

2.1 Introduction . . . 30
2.2 Related literature . . . 33
2.3 Sex ratios and historical context . . . 35
2.3.1 Sex ratios constructed from German Census data . . . 35
2.3.2 Variation in sex ratios . . . 36
2.4 Individual level data . . . 42
2.4.1 The German Socio-Economic Panel . . . 42
2.4.2 1 Percent sample of West German population . . . 45
2.5 Empirical strategy . . . 46
2.6 Results . . . 49

2.6.1 Life cycle patterns of fertility and marriage . . . 50
2.6.2 Analysis of mechanisms: out-of-wedlock births and matching quality 55
2.6.3 Robustness . . . 59
2.6.4 Results for East Germany . . . 62
2.7 Conclusions . . . 64

Appendices **65**
2.A Main Appendix . . . 65
2.B Online Appendix . . . 71
2.B.1 Regression results for robustness checks . . . 71
2.B.2 Construction of the data set on sex ratios across Census years . . . 75
2.B.3 References Census data . . . 79

3 Historic sex-ratio imbalances predict female participation in the market for politicians **83**
3.1 Introduction . . . 84
3.2 Related literature . . . 86
3.3 Data . . . 88
3.3.1 Outcome variables . . . 88
3.3.2 Sex ratios . . . 89
3.3.3 Mapping of county-level data into voting-district data . . . 93
3.3.4 Summary statistics . . . 93
3.4 Historical background and preliminary data analysis . . . 94
3.5 Results . . . 100
3.5.1 The sex ratio effect on female participation . . . 100
3.5.2 Candidate supply vs. voter demand effects . . . 103
3.6 Channels for long-lasting sex-ratio effects . . . 106
3.7 Robustness checks . . . 109
3.8 Conclusions . . . 116

Appendices **118**
3.A Data appendix . . . 118
3.A.1 Census data . . . 118
3.A.2 Election data . . . 123

4 Do local public expenditures on sports facilities affect sports participation in Germany? **124**
4.1 Introduction . . . 125
4.2 Data . . . 128

4.2.1 Regional expenditure data . . . 128
4.2.2 Individual level data: The German Socio-Economic Panel . . . 131
4.2.3 Descriptive statistics . . . 133
4.3 Econometrics . . . 137
4.3.1 The causal framework for continuous treatment . . . 137
4.3.2 Empirical strategy . . . 138
4.3.3 Estimation of the GPS . . . 140
4.3.4 Common support and balancing tests . . . 141
4.3.5 Estimation of the dose-response function . . . 145
4.4 Results . . . 145
4.4.1 Estimation of the GPS, common support and balancing tests . . . 146
4.4.2 Effects of sports expenditures on sports participation . . . 149
4.4.3 Robustness . . . 151
4.5 Conclusion . . . 152

Appendices **154**
4.A Sports expenditure data . . . 154
4.B Construction of SOEP sample and descriptive statistics . . . 155
4.C GPS-estimation, common support & balancing tests . . . 157
4.D Robustness checks . . . 161
4.E Additional estimation results . . . 166

5 Individual labor market effects of local public expenditures on sports **208**
5.1 Introduction . . . 209
5.2 Sports related public expenditures: data and measure . . . 211
5.3 Empirical strategy . . . 214
5.3.1 Sampling . . . 215
5.3.2 Identification . . . 216
5.3.3 Estimation . . . 219
5.4 Results . . . 220
5.4.1 Selection process . . . 221
5.4.2 Labor market effects . . . 223
5.4.3 Robustness checks . . . 227
5.4.4 Effect heterogeneity . . . 229
5.4.5 Exploring mechanisms . . . 231
5.5 Conclusion . . . 233

Appendices **236**
5.A Data cleaning, transforming, and enriching process 236
5.B Variable descriptions and further estimation results 243
5.C Robustness check 250
5.D Effect heterogeneity 254
5.E Exploring the channels 256

Bibliography **259**

Curriculum Vitae **281**

List of Tables

1.1 Affectedness by the reform by cohort . . . 7
1.2 Completed vocational training by sector - cohorts 1921/22 and 1923/24 - Census 1970 . . . 9
1.3 Average length of vocational training (in years) by cohort - Census 1970. . . 9
1.4 Socio-economic and social status by (Handl, 1977) and (Mayer, 1977) . . . 13
1.5 Summary Statistics Census 1970. . . . 14
1.6 Summary Statistics Microcensus 1971 . . . 15
1.7 Results social status husband - Census 1970 / MZU71 . . . 16
1.8 Results probability marrying down - Census 1970 / MZU71 . . . 18
1.9 Labor market outcomes 1970 - Census 1970 . . . 20
1.10 Results reduced form fictitious reforms - MZU71 . . . 22
1.11 Results social status husband - Handl score - Census 1970 / MZU71 . . . 23
1.12 Results probability marrying down - Handl score - Census 1970 / MZU71 . 23
1.13 Reduced form estimates marriage outcomes - Census 1970. . . . 24
1.14 Social status women before marriage - Census 1970 / MZU71 . . . 26
1.15 Labor force participation around the time of marriage - MZU71 . . . 27

2.1 Comparison of analytic sample to sample of West German women: outcomes and control variables . . . 44
2.2 Descriptive statistics 1% percent sample German Census 1970 . . . 46
2.3 Fraction men married to women between 2 years older and up to 6 years younger, Census 1950 . . . 48
2.4 Regression results of the impact of imbalances in sex ratios on fertility and marriage outcomes . . . 53
2.5 OLS regressions of imbalances in sex ratios on age at first marriage, age at first birth and out-of-wedlock births . . . 55
2.6 Matching by education, OLS regression of wife's years of education on husband's years of education . . . 57
2.7 OLS regression of spousal age gap on imbalances in sex ratios . . . 58
2.A.1 Regression results on the impact of sex ratios in 1946 and 1950 on women's age at birth of the first kid and the last kid . . . 65
2.A.2 Regression results on the impact of sex ratios in 1946 on fertility and marriage outcomes by women's age at birth, younger versus older ages . . . 66

2.B.1 Regression results of the impact of imbalances in sex ratios on total fertility – Table for Figure 2.A.1, robustness checks I 71
2.B.2 Regression results of the impact of imbalances in sex ratios on intensive margin of fertility – Table for Figure 2.A.2, robustness checks I 72
2.B.3 Regression results of the impact of imbalances in sex ratios on total fertility – Table for Figure A.3, robustness checks II 73
2.B.4 Regression results of the impact of imbalances in sex ratios on intensive margin of fertility – Table for Figure A.4, robustness checks II 74
2.B.5 States of Germany today and corresponding states in 1939 78

3.1 Descriptive statistics: district and candidate level 94
3.2 OLS regression of the share of female nominees in a constituency on 1946 sex ratios 102
3.3 Probit regression of the probability that candidate is female on 1946 sex ratio. Columns 1-2: candidate nominated for first vote, errors are clustered at the voting district level. Columns 3-4: candidate elected by first vote. . 105
3.4 OLS regression of 1946 sex ratio effect on 1997 female labor force participation.. 108
3.5 OLS regression of the share of female nominees in a constituency in the 1994 Bundestag elections on 1946 sex ratios. Sample includes Bavarian constituencies. 110
3.6 OLS regression of the share of female nominees in a constituency in the 1994 Bundestag elections on 1946 sex ratios. Sample excludes Bavarian constituencies. 110
3.7 Probit regression of the probability that candidate is female in the 1994 Bundestag elections on 1946 sex ratio. Columns 1-2: candidate nominated for first vote, errors are clustered at the voting district level. Columns 3-4: candidate elected by first vote. 111
3.8 OLS regression of the share of female nominees in a constituency in the 1987 Bundestag elections on 1946 sex ratios. 112
3.9 Probit regressions of the probability that candidate, affiliated with a right (CDU/CSU/FDP) or left (SPD/PDS/Grüne) wing party, is female on 1946 sex ratio.. 113
3.10 OLS regression of the share of female nominees in a constituency on 1946 sex ratios: West- and East Germany.. 114
3.11 Probit regression of the probability that elected candidate is female on 1946 sex ratio. Sample is restricted to candidates born in 1928 or after. 115
3.12 OLS regression of the mean population age in 1995 on 1946 sex ratio.. . . . 115

4.1 Aggregated community expenditures of municipalities in SOEP sample by year . . . 134
4.2 Descriptive Statistics (selected individual and regional characteristics) . . 136
4.3 Common support procedure . . . 144
4.4 Balancing before GPS adjustment . . . 148
4.5 Balancing after step 2 of common support procedure . . . 148
4.B.1 Descriptive Statistics . . . 156
4.C.1 Estimation of GPS . . . 157
4.C.2 Balancing before GPS adjustment . . . 160
4.E.1 Estimation of GPS - A1 sample (17-30 years) . . . 167
4.E.2 Estimation of GPS - A3 sample (46-65 years) . . . 168
4.E.3 Estimation of GPS - A1 sample with lagged outcome & treatment . . . 169
4.E.4 Estimation of GPS - A2 sample with lagged outcome & treatment . . . 170
4.E.5 Estimation of GPS - A3 sample with lagged outcome & treatment . . . 171
4.E.6 Estimation of GPS - A2 sample - small municipalities . . . 172
4.E.7 Estimation of GPS - A2 sample - large municipalities . . . 173
4.E.8 Balancing before GPS adjustment - A1 sample (17-30 years of age) . . . 195
4.E.9 Balancing after step 2 of common support procedure - A1 sample . . . 196
4.E.10 Balancing before GPS adjustment - A3 sample (46-65 years of age) . . . 197
4.E.11 Balancing after step 2 of common support procedure - A3 sample . . . 198
4.E.12 Balancing after step 2 of common support procedure - A1 sample with lagged treatment and outcome . . . 199
4.E.13 Balancing after step 2 of common support procedure - A2 sample with lagged treatment and outcome . . . 200
4.E.14 Balancing after step 2 of common support procedure - A3 sample with lagged treatment and outcome . . . 201
4.E.15 Balancing before GPS adjustment - A2 sample - small municipalities . . . 202
4.E.16 Balancing after step 2 of common support procedure - A2 sample - small municipalities . . . 203
4.E.17 Balancing before GPS adjustment - A2 sample - large municipalities . . . 204
4.E.18 Balancing after step 2 of common support procedure - A2 sample - large municipalities . . . 205

5.1 Mean values and marginal effects of the selection models (selected variables) . . . 222
5.A.1 Development of the number of municipalities between 2002 and 2006 in the RDB . . . 237

5.A.2 FIFA World Cup 2006 related expenditures of the 12 cities hosting games between 2001 and 2006 in EUR million 241
5.B.1 Variable description 243
5.B.2 Mean values and marginal effects of the selection models 244
5.B.3 Common support statistics for the estimated matching models for men (with / without lagged SPE) and women (without lagged SPE) 245
5.B.4 Balancing statistics with the mean and median biases for the estimated matching models for men. 246
5.B.5 Balancing statistics with the mean and median biases for the estimated matching models for men with lagged SPE. 247
5.B.6 Balancing statistics with the mean and median biases for the estimated matching models for women. 249
5.E.1 Average values of the SPE measures and different indicators of sports facility provision in the 16 states (Länder) in Germany 257
5.E.2 Comparing the effects for men between our main specification (ATE) and the main specification including various health measures from the corresponding year as additional covariates in the selection process (ATE_{Health}). 258

List of Figures

1.1 Share of women with completed vocational training by cohorts and year of birth - Census 1970 8
1.2 Social status of husbands for women born in 1922 versus women born in 1924. 17
1.3 Categories of net earnings of cohorts 1921/22 and 1923/24 21
1.4 Labor force status around marriage across birth cohorts. 26

2.1 Sex ratios by year of birth in 1939 and in 1946 37
2.2 Share of male birth cohort that died in war by year of birth 38
2.3 Sex ratios of overall population and native population on state level in 1946 40
2.4 Sex ratios in 1950, 1961 and 1970, West Germany 41
2.5 Persistency in sex ratios over the life cycle for birth cohorts 1924-1932, state of Bavaria 42
2.6 Comparison of sex ratios using +3/-2 birth cohorts to sex ratios with men and women from same birth cohorts 49
2.7 Coefficients and 95% CI for the impact of unbalanced sex ratios in 1946 on total fertility and marriage over the life cycle 52
2.8 Distribution of total fertility for women with low sex ratios (born before 1927) and balanced sex ratios (born after 1927) 54
2.9 Coefficients and 95% CI for the impact of unbalanced sex ratios in 1946 on the extensive and intensive margins of fertility over the life cycle in East Germany. 63
2.A.1 Regression results of the impact of imbalances in sex ratios on total fertility - Robustness checks I 67
2.A.2 Regression results of the impact of imbalances in sex ratios on the intensive margin of fertility - Robustness checks I 68
2.A.3 Regression results of the impact of imbalances in sex ratios on total fertility - Robustness checks II 69
2.A.4 Regression results of the impact of imbalances in sex ratios on the intensive margin of fertility – Robustness checks II 70

3.1 Pre- and post-war county level sex ratio variation in Germany 90
3.2 Sex ratios in 1946, 1950 and 1964 by state 92
3.3 Sex ratios in 1946, 1950, 1961 and 1970 by state 93

3.4 Sex ratios in 1939 by post-unification state and year of birth. 96
3.5 Sex ratios in 1946 by post-unification state and year of birth. 97
3.6 Sex ratios in 1946 of overall population compared to sex ratios of native population by post-unification state and year 99
3.7 Sex ratio densities for constituencies with female/male winners. 103

4.1 Distance weighted per capita sports expenditures 131
4.2 Timeline of measuring information 132
4.3 Sports expenditures full range vs. restricted to positive values up to the 95th percentile 134
4.4 Common support 147
4.5 Results main specification 150
4.C.1 Common support - A2 sample (main sample) 158
4.D.1 Results different specifications outcome variable 161
4.D.2 Results A1 sample (17-30 years of age) 162
4.D.3 Results A3 sample (46-65 years of age) 163
4.D.4 Results with lagged outcome and treatment 164
4.D.5 Results small / large municipalities 165
4.E.1 Common support - A1 sample (17-30 years) 174
4.E.2 Common support - A3 sample (46-65 years) 177
4.E.3 Common support - A1 with lagged outcome and treatment (17-30 years) . 180
4.E.4 Common support - A2 with lagged outcome and treatment (31-45 years) . 183
4.E.5 Common support - A3 with lagged outcome and treatment (46-65 years) . 186
4.E.6 Common support - A2 sample - small municipalities 189
4.E.7 Common support - A2 sample - large municipalities 192
4.E.8 Additional results A1 sample (17-30 years of age) 206
4.E.9 Additional results A3 sample (46-65 years of age) 207

5.1 Distance weighted sports facility per capita expenditures in Germany. . . . 213
5.2 Effects of different levels of SPE on household income and earnings. . . . 224
5.3 Effects of different levels of SPE on working time, wages and employment status. 226
5.A.1 Distribution of geometric centers (centroids) for each municipality in Germany 242
5.C.1 Effects of different levels of SPE on household income and earnings during the exposure period including (main specification, black lines) / excluding (red lines) any individuals who were still in the education system in 2001-2006. 250

5.C.2 Effects of different levels of SPE on working time, wages and employment status during the exposure period including (main specification, black lines) / excluding (red lines) any individuals who were still in the education system in 2001-2006. 251
5.C.3 Effects of different levels of SPE on household income, earnings, working time, wages and full time employment for men when including lagged SPE as additional control. 252
5.C.4 Effects of different levels of SPE on household income, earnings, working time, wages and full time employment for men when using the pure SPE measure. 253
5.D.1 Effects of different levels of SPE on earnings for men 20-36 / 37-52 years old in 2001.. 254
5.D.2 Effects of different levels of SPE on working time, wage and employment status for men 20-36 / 37-52 years old in 2001. 255
5.E.1 Effects of different levels of SPE on household income, earnings, working time, wages and full time employment for men during the post-exposure period excluding (main specification, black lines) / including (red lines) various health measures from the corresponding year as additional covariates in the selection process. 256

Summary

This thesis combines five essays in the areas of Labor and Sports Economics. While the individual parts address quite different research questions, all chapters aim to identify causal effects. The first chapter investigates effects of vocational training for women on partner choice and labor market outcomes in a society with traditional gender roles. Chapter 2 and 3 study effects of gender imbalances caused by World War II on fertility and labor markets in Germany. The last two chapters analyze the effect of local public sports expenditures on sports participation and individual labor market outcomes in Germany.

The first chapter finds that women who obtained vocational training married a partner with higher social status. When we compare the social status of a woman's father and spouse, we see that women were also less likely to marry down. Exposure to vocational training did not impact labor market outcomes though. The essay provides evidence that prevalent social norms at the time potentially played an influential role.

The second chapter studies the effect gender imbalance has on fertility over the life cycle. The study shows that results strongly depend on at what point in the life cycle effects are evaluated and that the decomposition of fertility into extensive and intensive margins is essential. We find that women affected by the gender imbalance marry later and are more likely to remain childless throughout life. However, they catch up and even overcompensate at later ages with regard to number of children conditional on having any children.

The third chapter investigates the effects of unbalanced gender ratios on female labor force participation, focusing specifically on women's participation in the political labor market. The analysis shows that in Germany's 1990 elections, women were more likely to run for office in voting districts where sex ratios were more unbalanced after World War II. However, gender imbalances did not affect the probability of a woman winning in the election, and, thus, it appears that voters were not more inclined to vote for women.

The fourth chapter combines local public expenditure data with individual level data and studies how the effects of expenditures on sports facilities change with different levels of sports expenditures. Since individuals also benefit from the expenditures of neighboring communities, we construct locally weighted averages of expenditures. The study finds that expenditures on sports infrastructure has no effect on the probability to engage in sports.

The fifth chapter finds that high expenditure on sports facilities impact male earnings positively while no effects are found for women. The study further shows that the observed patterns seem plausible, since women practice sports with a lower likelihood compared to men in general and are, in addition, less likely to be active in publicly funded sports infrastructure if they engage in sports.

Zusammenfassung

Die vorliegende Dissertation besteht aus fünf Aufsätzen im Bereich der Arbeitsmarkt- und Sportökonomie. Während die einzelnen Kapitel unterschiedliche Fragestellungen adressieren, haben alle Kapitel die Identifikation von kausalen Effekten zum Ziel. Das erste Kapitel analysiert Effekte von Berufsausbildung auf Partnerwahl und individuelle Arbeitsmarktergebnisse für Frauen in einer von traditionellen Rollenbildern geprägten Gesellschaft. Kapitel 2 und 3 analysieren Auswirkungen des Männermangels nach dem 2. Weltkrieg in Deutschland auf Fertilität und Arbeitsmärkte. Die letzten beiden Kapitel untersuchen den Effekt von öffentlichen Ausgaben für Sportstätten auf sportliche Aktivität sowie individuelle Arbeitsmarktergebnisse.

Das erste Kapitel zeigt, dass Frauen mit abgeschlossener Berufsausbildung einen Partner mit höherem sozialen Status heirateten. Beim Vergleich von sozialem Status der Ehemänner und Väter der Frauen sieht man zudem, dass diese auch mit geringerer Wahrscheinlichkeit "nach unten" heiraten. Individuelle Arbeitsmarktergebnisse wurden durch den Abschluss einer Berufsausbildung jedoch nicht beeinflusst. Der Aufsatz zeigt abschliessend auf, dass damals vorherrschende soziale Normen möglicherweise eine einflussreiche Rolle spielten.

Das zweite Kapitel untersucht den Effekt von Ungleichgewichten in der Anzahl an Männer zur Anzahl an Frauen auf Fertilität. Die Studie zeigt, dass gewonnene Erkenntnisse stark davon abhängen, zu welchem Alter der Frauen man Fertilität betrachtet. Darüber hinaus ist die Dekomposition von Fertilität in Kinderlosigkeit (*"extensive margin of fertility"*) sowie Kinderzahl, wenn sie Kinder bekommen (*"intensive margin of fertility"*), von grosser Bedeutung. Die Analysen zeigen, dass die vom Männermangel betroffenen Frauen später heiraten und mit höherer Wahrscheinlichkeit kinderlos bleiben. Jedoch holen sie bezüglich Anzahl an Kindern auf und überkompensieren sogar in späteren Lebensjahren, wenn sie Kinder bekommen.

Das dritte Kapitel betrachtet Effekte des Männermangels auf die Arbeitsmarktbeteiligung von Frauen und fokussiert sich auf Politiker. Die Studie zeigt, dass bei den Bundestagswahlen 1990 Frauen mit grösserer Wahrscheinlichkeit in Wahlkreisen für ein Mandat kandidierten, welche nach dem 2. Weltkrieg einen grösseren Männermangel aufwiesen. Jedoch ist kein Effekt vom Geschlechterverhältnis auf die Wahrscheinlichkeit, dass eine Frau die Wahl gewinnt, beobachtbar. Daraus lässt schliessen, dass Wähler nicht stärker dazu geneigt waren eine Frau zu wählen.

Das vierte Kapitel kombiniert öffentliche Ausgabedaten auf kommunaler Ebene mit Individualdaten und analysiert, wie sich Effekte von Ausgaben für öffentliche Sportstätten mit unterschiedlicher Ausgabenhöhe verändern. Da Individuen auch von Ausgaben benachbarter Gemeinden profitieren, konstruieren wir lokal gewichtete Durchschnitte von

Sportausgaben. Die Studie kommt zum Ergebnis, dass öffentliche Ausgaben für Sportstätten keinen Effekt auf die Wahrscheinlichkeit haben sportlich aktiv zu sein.

Das fünfte Kapitel zeigt, dass hohe Sportstättenausgaben mit positiven Einkommenseffekten für Männer einhergehen, während für Frauen keine Effekte beobachtet werden. Im Weiteren zeigt die Studie, dass die beobachteten Muster plausibel erscheinen, da Frauen im Allgemeinen weniger wahrscheinlich sportlich aktiv sind als Männer. Darüber hinaus üben sie Sport mit einer geringeren Wahrscheinlichkeit in öffentlich geförderten Sportstätten aus, wenn sie sportlich aktiv sind.

1 Effects of vocational training for women in a society with traditional role models – Evidence from West Germany[1]

Abstract

This paper studies the effect of vocational training on partner choice and labor market outcomes for women in a society with traditional role models. To this end, we exploit a German labor market law which strongly increased the supply of available spots for vocational training. Applying a two-sample IV approach, we find that women who completed vocational training married a partner with higher social status. They were also less likely to marry down when comparing social status of their father and husband. Female labor market outcomes in later life are however not affected by the exposure to vocational training. We provide some evidence that prevailing social norms might have played a strong role.

Keywords: Vocational training; Marriage market; Female labor force participation; Social norms

JEL codes: J12; J16; J24; N34

[1]I am very grateful to Michael Lechner, Michael Knaus, Michael Zimmert, Iris Kesternich, Bettina Siflinger and Joachim Winter for helpful comments and to Bernhard Schimpl-Neimanns for providing expertise on the 1970 Census and the Microcensus supplementary survey of 1971. All remaining errors are my own. The author gratefully acknowledges financial support by the Swiss National Science Foundation through grant SNF 100018E-161598.

1.1 Introduction

In German speaking countries, vocational training has traditionally been an integral part of the educational system and has been repeatedly called the backbone of the economy. In recent years, other countries showed a growing interest in the system of vocational training. In 2012, Greece, Italy, Latvia, Portugal, and Slovakia engaged in cooperation with Germany to develop and implement a system of vocational education and training in their countries based on the German model (European Centre for the Development of Vocational Training, 2017). Previous evaluations of returns to vocational education have often focused on the male population, in part, on the ground that male participation on the labor market was more stable over the life-cycle than female participation in the past.[2] Indeed, while total number of years worked have remained largely the same for men over the last 50 years,[3] changes are drastic for the female population. While women in West Germany aged 60 to 65 in 1966 had worked on average for 20 years (Kirner, 1980, p. 30), women between 65 and 75 years of age in 2015 pursued employment in total for around 30 years - and the trend is rising (TNS Infratest, 2017).

In this paper, we study effects of vocational training for women in West Germany. The former West Germany was a conservative society with traditional role models and social norms, where the expected role for married women was to take care of the household and children (see e.g. Moeller, 1993).[4]

Understanding the role of social norms in the analysis of effects of education for women and in female labor force participation is important beyond a region specific context. For instance, Fernández (2007) finds that for female immigrants the culture of their parents home country is still an important determinant of their labor force participation in the US. While female labor force participation and attitudes towards working women have changed and evolved considerably over the course of the 20th century (Fernández, 2013; Hazan & Maoz, 2002), there still exist regions with low levels of female labor force participation today, in particular in the Middle East and North Africa (World Bank, 2018). Culture

[2] See for instance Hanushek, Schwerdt, Woessmann, and Zhang (2017).

[3] Total number of years worked were on average 43 years in the 1960s (Kirner, 1980, p. 30) and 42 years in 2015 (TNS Infratest, 2017) for men.

[4] The term for the traditional family model in West Germany was "Hausfrauenehe" (housewife family model). The conservative nature of the West German society was also reflected in the fact that until 1977, women in West Germany needed their husband's consent to work (see e.g. Schulz, 1998, p. 130).

and social norms have been found to play an important role in this (Bursztyn, González, & Yanagizawa-Drott, 2018; Göksel, 2013). With increased immigrant related diversity, it is important for host countries to understand which social norms are influential in their migrant communities.

Since there is typically selection into pursuing vocational education, we exploit a labor market law which strongly increased the supply of vocational training spots in Germany in 1937 from one school leaving cohort to the next. We use affectedness of cohorts by the reform as an instrument for vocational training. Throughout the analysis, we focus on women who graduated from the basic school track (87% of women at the time). After graduation, they typically either obtained vocational education or began to work as an unpaid family member or unskilled worker.

In a first step, we analyze whether vocational training impacted marriage market outcomes in terms of social status of the husband. Since three years of vocational training in a firm and school quite likely broadens the set of potential partners, we are interested in whether vocational training changed social opportunities through marriage. Since a woman's professional career often ended with marriage or with birth of the first child (Niehuss, 2001), the occupational status of the husband defined essentially the social status of the wife.[5] This implied that at least maintaining social status relative to their family background was important for women. We therefore also compare social status of the husband and father and study whether vocational education impacted the likelihood of marrying down. Additionally, we check whether the marriage rate, age at marriage and number of children are altered.

In a next step, we look at labor market outcomes in later life, when women in our sample are between 45 and 48 years of age. From an incentive theory perspective, one would expect that women with more education are more active on the labor market. Having said this, taking into account traditional role models and prevailing social norms at the time, it appears unclear whether one would expect that completing vocational education should affect labor market outcomes for women.

While we can use a 1% sample of the West German Census in 1970 to estimate effects

[5] 75.9% of women in our sample were working right before marriage. After marriage, 39.6% of women were still in the labor force. In 1970, where these women were between 45 and 48 years of age, 46.8% were working.

of vocational education on labor market outcomes, no adequate information on partners and parents is available in the data set. This information, in turn, is available in the Microcensus Supplementary Survey from 1971 (MZU 71 in the following). Since year of birth information is available in both data sets, we can apply a two-sample instrumental variable approach to estimate the effect of vocational education on social status of the husband. The appeal of the estimation approach comes from being able to estimate the first stage and reduced form in two separate data sets under the condition that both datasets must stem from the same underlying population.[6]

We find that women marry a husband of higher social status. They are also significantly less likely to marry down when comparing social status of their father and husband. Vocational training does however not affect female labor market outcomes measured in 1970. Also, the marriage rate, age at marriage and number of children are not altered.

We present some evidence that social norms might play a strong role by looking at decisions around the time of marriage. Marriage was the point in life where the adequate role of a woman is expected to change drastically. We look at occupational status right before marriage and whether women were more likely to stay in the labor market after marriage. We find while women with vocational training are more likely to have a job position with higher social status before marriage, this does not increase the likelihood that they continue working after they get married. This indicates that social norms might play a strong role.

The remainder of this paper is structured as follows. Related literature is reviewed in Section 1.2. Section 1.3 provides a detailed discussion about the institutional background and the reform. Section 1.4 discusses identification as well as the estimation strategy and Section 1.5 describes the data. Section 1.6 presents the main results. Robustness checks and a potential channel are discussed in Section 1.7. Section 1.8 concludes.

[6]Influential early papers on two-sample estimation are Klevmarken (1982), Angrist and Krueger (1992, 1995) and Arellano and Meghir (1992). Björklund and Jäntti (1997), Dee and Evans (2003) and Devereux and Hart (2010) are examples of applications.

1.2 Related literature

This paper is related to the literature evaluating returns to vocational education (see e.g. Fersterer, Pischke, and Winter-Ebmer (2008); Meer (2007); Oosterbeek and Webbink (2007) for examples of recent studies). Golsteyn and Stenberg (2017) and Hanushek et al. (2017) take a life cycle perspective in their comparison of returns to vocational education and general education. They find a trade-off between a relative advantage of vocational education over general education in the phase of labor market entry and a relative disadvantage in later life. Malamud and Pop-Eleches (2010) make use of an educational reform in Romania which shifted people from vocational training to general education and find that men affected by the policy display no significant differences in earnings and labor force participation compared to men not affected by the reform. All in all, with the exemption of Golsteyn and Stenberg (2017) and Meer (2007), these studies focus on the male population. Therefore, little is known about the effects of vocational education for women and we aim at adding to the literature in this dimension.

Secondly, this paper relates to research studying the marriage market and how matches are formed on this market. According to Becker (1973), people get married if both spouses reach a higher expected utility through marriage compared to staying unmarried. With regards to spousal selection, it is well established in the literature that mating tends to be rather positive assortative which means that the characteristics of spouses display positive correlation (Becker, 1991; Hitsch, Hortaçsu, & Ariely, 2010). Similarity of partners in terms of education, attractiveness or height have been documented in previous research (e.g. Becker, 1973; Browning, Chiappori, & Weiss, 2008; Kalmijn, 1998; Oreffice & Quintana-Domeque, 2010). Independent of individual preferences, sorting might also reflect natural meeting points such as schools or the workplace where people spend much time and which are socially structured (Browning et al., 2008; Hitsch et al., 2010). We contribute to this literature by providing evidence on how exposure to vocational training affects partner choice in a society with traditional role models.

1.3 Vocational training requirement in Germany

The Nazi administration required firms to offer apprenticeships proportional to the number of skilled workers ("Facharbeiter") in a firm in a wide range of sectors starting April, 1937. This requirement was part of the Four Year Plan ("Vierjahresplan") and aimed at securing the supply of skilled workers.[7] It was enacted since the Nazis feared a shortage of skilled worker in sectors of military importance (Wolsing, 1977, p. 162). The metal, iron as well as the construction industry were affected by this ordinance.[8] Firms who did not comply with the rule had to pay a fee per vocational training spot they did not supply (Mason, 1975, p. 224).[9] Consequences of the reform were twofold. It increased the overall number of spots for vocational training and redirected young men towards these sectors. This in consequence opened up vocational training places for young women (in particular commercial training), since men were largely absorbed in the metal sector. In 1934/1935, there were 84,300 apprentice positions for women, 221,111 in 1938/1939 and 357,000 in 1942/1943 (Gillingham, 1986).[10]

Young people began vocational training after they finished school. Training typically lasted three years, took place in firms and was combined with days at the vocational school ("Berufsschule"), where theoretical concepts were taught. We look at women who graduated from the basic school track ("Volksschule"). At the time, attending the basic schooling track was very common for young women (around 87% of women in a cohort). Only a small share of women obtained a higher educational degree back then.

The enactment of the law created a natural experiment with the first cohort graduating from school after April 1937 being exposed to an increased supply of apprenticeship positions. In the following, we outline the institutional set-up of the schooling system at

[7] The German title was "Erste Anordnung zur Durchführung des Vierjahresplans über die Sicherstellung des Facharbeiternachwuchses vom 7. November 1936" and the original text of the reform can be found in Mason (1975, p. 223ff).

[8] More specifically, the ordinance encompassed the following sectors: Building construction, civil engineering and road construction. Within the metal and iron industry the following industries were affected: iron and steel producing, non-ferrous metal, foundry, mechanical engineering automotive, aviation, electrical, steal and iron, precision mechanics and the optical industry (Wolsing, 1977, pp. 163-164).

[9] This fee was 50 "Reichsmark" per trainee per month (Mason, 1975, p. 224). In 1936, weekly gross earnings of a skilled (unskilled) worker in the metal sector were on average 49.28 Reichsmark (32.96 Reichsmark) (Statistisches Reichsamt, 1937, p. 315).

[10] Exact numbers by school leaving cohort are unfortunately not readily available. Gillingham (1986) notes that around two third of apprentice positions for girls were in the secretarial field or other office positions.

the time and identify cohorts which were affected by the reform. Throughout the 1920s and the 1930s, the school year ended around Easter and children started school at the age of six. Since we have exact date of birth information in the census data, we can construct school cohorts, where one cohort encompasses women born after July 1 until June 30th of the following year.[11] There were eight years of compulsory schooling, however rural regions in the south (most parts of Bavaria and of the later state Baden–Württemberg) had only seven years of schooling.[12]

As shown in Table 1.1, cohort 1923/24 was the first cohort which was fully affected by the reform, cohort 1921/22 was the last cohort not affected. Different lengths of compulsory schooling within Germany involve that only a part of cohort 1922/23 is affected by the reform.

Table 1.1: Affectedness by the reform by cohort

cohort	Start school	End of schooling	End of school. South	affected by reform
1921/22	Easter 1928	Easter 1936	Easter 1935	No
1922/23	Easter 1929	Easter 1937	Easter 1936	Not in south
1923/24	Easter 1930	Easter 1938	Easter 1937	Yes

Individuals of this cohort were not exposed to the reform if they graduated from school in the south of Germany. Individuals of this cohort who finished school in all other parts of Germany were in contrast exposed to the reform. However, since the census does not contain information about residence during childhood, we cannot infer whether a woman from cohort 1922/23 lived in the south when graduating from school and was hence not exposed to the reform. We hence exclude the cohort 1922/23 from the analysis.

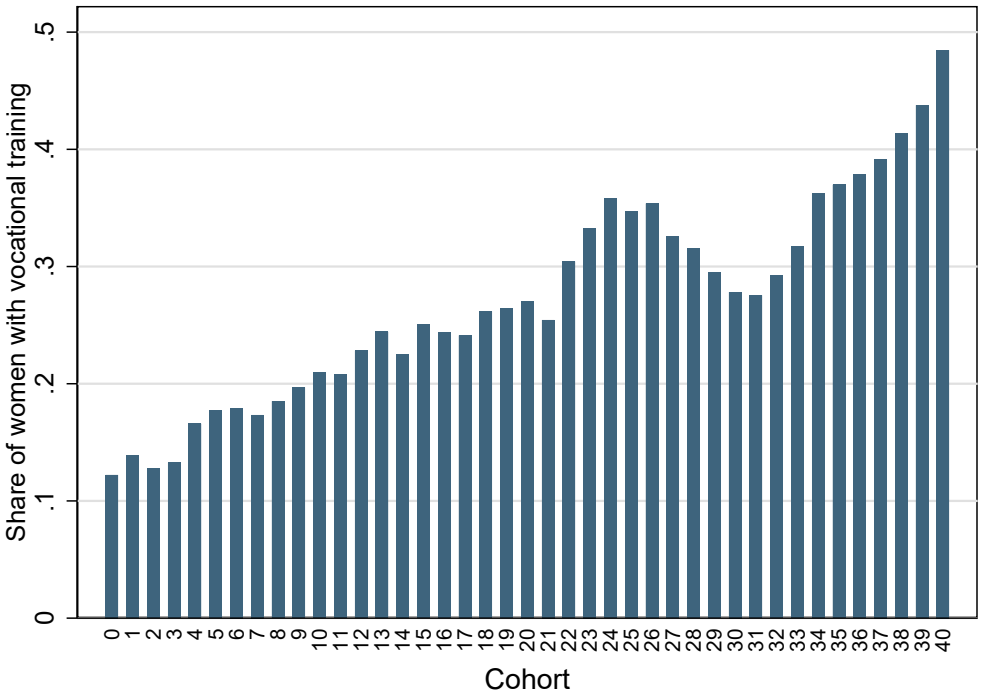
Figure 1.1a displays the share of women obtaining vocational training by school cohort for cohorts 1900/01 to 1940/41 using data from the 1970 census. Increasing shares of women completed vocational training over the 40 years time frame with a distinct jump from cohort 1921/22 to 1922/23. Figure 1.1b shows attainment of vocational training by year of birth from 1900 to 1940.[13]

[11] The effective data for school enrolment was June 30 in all parts in the Free State of Prussia, the by far largest German state during the Weimar Republic (Faust, 2006).

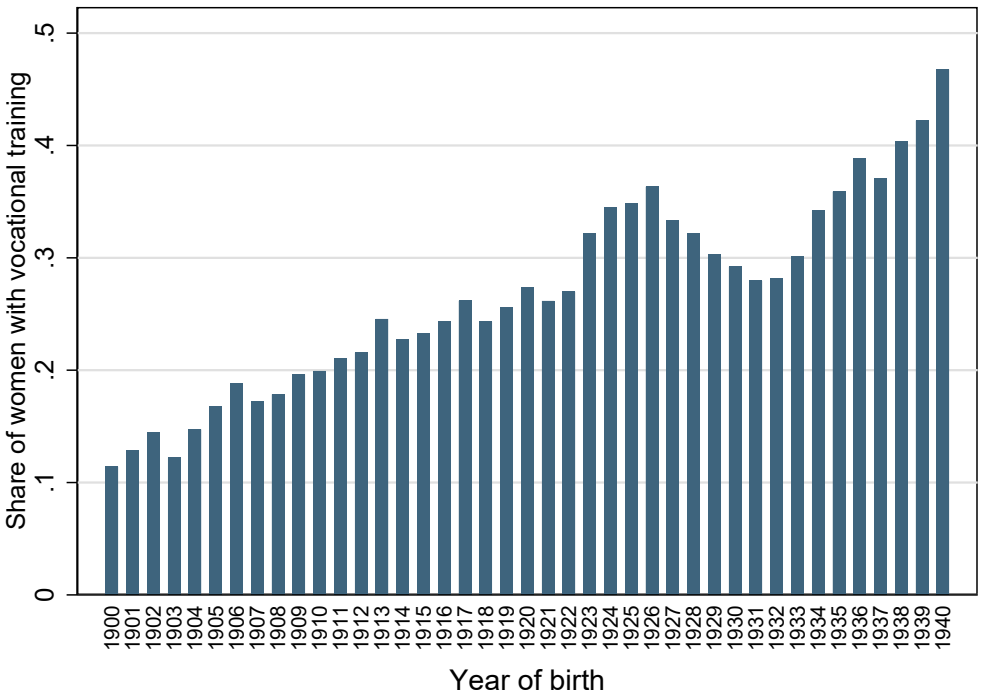
[12] The eighth year of compulsory education was not introduced in all regions of Bavaria until 1938 (Fenn & Körner, 2007, p. 414).

[13] We also show completion of vocational training by year of birth, since for parts of the analysis we can only work with year of birth information. This will be outlined in Section 1.4.

Figure 1.1: Share of women with completed vocational training by cohorts and year of birth - Census 1970



(a) Share of women obtaining vocational training by school cohorts



(b) Share of women obtaining vocational training by year of birth

Table 1.2 illustrates from which sectors this overall increase in vocational education originates from by comparing cohorts 1921/1922 and 1923/1924. Completed apprenticeships increased particularly in the area of sales of goods (1.8 percentage points) and office jobs (2.1 percentage points) and to a smaller extent textile, health and services in general. The average length of vocational training was not altered by the requirement as can be seen in Table 1.3.

Table 1.2: Completed vocational training by sector - cohorts 1921/22 and 1923/24 - Census 1970

Sector vocational training	Cohort		
	1921/22	1923/24	Both cohorts
	%	%	%
No vocational training	74.5	66.7	70.8
Agriculture, animals	0.6	0.7	0.7
Material, Pottery, Glass, Chemistry, synthetic, paper, print, wood, metal, fitter, electrician	0.3	0.5	0.4
Textile	3.9	5.2	4.5
Leather	0.9	1.0	0.9
Mason, carpenter, road construction, interior designer, painter, machinist	0.0	0.1	0.1
Technical professions	0.2	0.3	0.3
Sales Goods	6.6	8.4	7.5
Sales Services	0.3	0.4	0.3
Traffic	0.1	0.2	0.2
Office jobs	8.2	10.3	9.2
Security	0.0	0.1	0.0
Writing, arts	0.1	0.1	0.1
Health	1.3	1.9	1.6
Social, education	0.8	1.0	0.9
Services	2.3	3.0	2.6
Total	100.0	100.0	100.0

Table 1.3: Average length of vocational training (in years) by cohort - Census 1970

Cohort	mean duration vocational training	Share completing vocational training
1920/21	2.786	0.270
1921/22	2.789	0.254
1922/23	2.780	0.304
1923/24	2.788	0.333
1924/25	2.727	0.358

1.4 Identification and estimation strategy

We estimate the effect of interest with an instrumental variable approach, where we use affectedness of the cohorts by the reform as an instrument for vocational education with $Z_i = 1$ for cohort 1923/24 and $Z_i = 0$ for cohort 1921/22.[14] As usual, while relevance of an instrument can be tested (Section 1.6.1), the assumption that the instrument influences the outcome only through its impact on the treatment (exclusion restriction) is not testable and we will discuss its plausibility in the following.

Both cohorts were born during the Weimar Republic. Hence, they were too young to be directly affected by events surrounding World War I. Important to our application, male cohorts to both sides of the threshold were fully drafted to the German armed forces during World War II.[15] This will be important when we look at social status of the husband, since it would be a major threat to identification if women in the cohort before the cut-off married men who were drafted and the ones born afterwards not (or vice-versa). The cohorts are between 45 and 48 years of age when labor market outcomes are evaluated. At this point in the life-cycle, the childbearing phase is most likely completed and hence should not impact labor market participation.[16]

Finally, we need to assume monotonicity which rules out so called "defiers" (Angrist, Imbens, & Rubin, 1996). In our case, a defier would be a woman who would complete vocational training if she was born in cohort 1921/22, but would not obtain vocational education if she was born in cohort 1923/24.

If the set of assumptions is fulfilled, we can identify the local average treatment effect (LATE). This gives us the average effect for the group of women whose treatment status is changed by the instrument. Since we have a binary instrument as well as a binary treatment variable, we can estimate our main specification with a Wald estimator:

[14] The overall set-up has resemblance to a fuzzy regression discontinuity design, where a local IV in a neighbourhood of an threshold x_0 is estimated (Hahn, Todd, & Van der Klaauw, 2001; Imbens & Lemieux, 2008). However, due to different lengths of compulsory schooling across Germany at the time, we have to exclude cohort 1922/23 from the analysis as discussed beforehand. In consequence, the set-up is distinct from a standard Fuzzy RDD set-up.

[15] Male cohorts from 1910-1927 were fully drafted (Overmans, 2004).

[16] 2% of women born in cohort 1921/1922 live in households with children below the age of 3. This is the case for 2.3% of women in the cohort 1923/24. The difference is not statistically significant.

$$\hat{\tau} = \frac{\hat{E}[Y_i|Z_i = 1] - \hat{E}[Y_i|Z_i = 0]}{\hat{E}[D_i|Z_i = 1] - \hat{E}[D_i|Z_i = 0]} \tag{1.1}$$

where $D_i = 1$ if a women obtained vocational training and $D_i = 0$ if she did not.

While we can estimate labor market outcomes with a "standard" Wald-estimator using census data of 1970, not all information we need to investigate whether vocational education influenced partner choice is available in the same data set. The census of 1970 has detailed information on vocational training (D_i) and date of birth (Z_i). It however lacks information on the partner for divorced or widowed women.[17] Also, no parental characteristics are available in the census. We need this information though, if we want to assess whether a woman married down relative to her family background. The MZU 71, on the other hand, provides information on the husband (Y_i) and also on birth cohorts (Z_i). Since the two data sets stem from the same population and the instrument is available in both data sets, we can apply a two-sample IV-approach where we estimate the first stage with the census data and the reduced form with the MZU 71 (Angrist & Krueger, 1992; Inoue & Solon, 2010). Standard errors are calculated using the Delta method (Devereux & Hart, 2010; Wassermann, 2004).

1.5 Data

Before we present both data sets in the following, we will briefly outline how our sample is defined. We apply the same sample restrictions to both data sets: Our sample contains women who lived already before the war within the later borders of West Germany. We hence exclude women who were expellees or refugees from the Soviet Zone. Furthermore, we exclude West-Berlin due to its special status and the state of Saarland (was not part of West Germany until 1957) from the analysis.

1.5.1 Census 1970

We use individual level data from the German census of 1970 to estimate the first stage as well as labor market outcomes for women. The dataset is a 1% sample of the West German population in 1970 and contains detailed information about vocational training

[17] 7.4% of women from cohorts 1919/1920 to 1925/1926 are widowed and 3.5% are divorced.

and also the exact date of birth.[18] Restricting the sample to the two cohorts 1921/22 and 1923/24 and applying the sample restrictions discussed above leaves us with a sample of 5,099 women.

1.5.2 Microcensus Supplementary Survey 1971 (MZU 71)

To investigate whether exposure to vocational training altered partner choice, we use data from the Microcensus Supplementary Survey *"Occupational and Social Shifts of the Population"* in 1971 which provides information on these two dimensions. As the census data from 1970, the Microcensus Supplementary Survey is an 1% sample of the German population and contains around 456,000 individuals (Tegtmeyer, 1979). The survey was specifically designed to provide information about the social as well as occupational situation of the population over a thirty year time frame and comprises retrospective information for the years 1939, 1950 and 1960. Compared to the census of 1970, the MZU 71 has the drawback of not containing the exact date of birth. Since only year of birth information is available, school cohorts cannot be constructed. To deal with this issue, we leave out the birth cohort of 1923 from the analysis of partner choice. The whole cohort of 1924 was exposed to the reform. People living in the former Prussian states who were born in the second half of 1922 were affected by the reform. Due to the lack of exact birth dates in the MZU 71, we however cannot increase precision and improve on that dimension. We hence underestimate the jump in the first stage, but also the reduced form. Hence, for this part of the analysis, $Z_i = 1$ if year of birth is 1924 and $Z_i = 0$ if year of birth is 1922.

Obviously, when investigating whether women exposed to vocational training married up in terms of social status of their husband, a key challenge is to measure a concept such as socio-economic or social status in an appropriate way.[19] Common international prestige scores such as ISEI or SIOPS, which would be typically used for this, cannot be constructed with older German Microcensus data. The reason for this is that occupational classification is not fine enough in these data sets. This being said, we can look at measures for social/socio-economic status which were constructed for data sets in the

[18] Schimpl-Neimanns and Frenzel (1995) provide a detailed description of the data set.

[19] It would be also interesting to look at income of the husband. However, income of the partner can only be constructed for married women. Also, income information is only available in very coarse intervals.

70s. The MZU 71 provides information on occupational status (26 different categories for the working population) as well as on economic sector. Mayer (1977) and Handl (1977) classified 34 categories based on occupational status and economic sectors and constructed metric prestige scores for these 34 categories.[20] Table 1.4 displays all categories and their respective scores.

Table 1.4: Socio-economic and social status by (Handl, 1977) and (Mayer, 1977)

Occupational status	Score Socio-economic status (Handl)	Score Social status (Mayer)
Farmer -10ha	35	43
Farmer -20ha	45	63
Farmer -50ha	72	100
Farmer +50ha	105	135
Free professions 1 employee	344	269
Free professions -49 employees	362	311
Self-employed -1 employee	157	132
Self-employed -9 employees	216	162
Self-employed +9 employees	284	208
Civil servant	73	99
Middle level civil servant	122	152
Upper-middle-level civil servants	179	215
Senior civil servants	344	278
Master workman	124	141
Employee	92	127
Employee civil service	77	122
Middle level employee	153	167
Middle level employee civil service	146	158
Upper-middle-level employee	271	207
Upper-middle-level employee civil service	147	196
Senior employee	310	247
Unskilled worker farming	10	10
Unskilled worker industry	23	34
Unskilled worker construction	19	12
Unskilled worker service sector	28	50
Semi-skilled worker industry	34	65
Semi-skilled worker construction	35	53
Semi-skilled worker service sector	39	76
Skilled worker farming	46	48
Skilled worker industry	56	97
Skilled worker construction	56	76
Skilled worker service sector	55	99
Assistant Foreman	70	101
Foreman	99	125

Scores can only be assigned for working population. *Source:* Mayer (1979, pp. 106-107).

[20] Mayer (1979) notes the high correlation of both scores. This is also the case for our sample, where the correlation of both scores is 0.93. This is despite both scores were constructed from different information and also the methodical approach was different. The score by Handl is based on empirical distribution of occupational status, education, house ownership and income. Spatial representation of marriage distances are the basis for the Mayer score.

1.5.3 Summary statistics

Table 1.5 provides descriptive statistics for our sample with the census data. 8.9% of women are not married and the average age of first marriage is 25.7. 78.4% of women have children and they are on average 26.4 years old when they get their first child. 29.2% obtain vocational education and in 1970, when they are around 45-48 years old, 47.2% of women are working.

Table 1.5: Summary Statistics Census 1970

	1921/22 & 1923/24			1922 & 1924		
Variable	N	Mean	SD	N	Mean	SD
Year of birth	5,099	1922.47	1.12	4,858	1922.96	1.00
Vocational education	5,098	0.292	0.455	4,857	0.306	0.461
Not married	5,099	0.089	0.284	4,858	0.089	0.285
Age at first marriage	4,160	25.67	4.86	3,986	25.49	4.80
Age at first child	3,996	26.43	4.74	3,783	26.22	4.66
Working 1970	5,099	0.472	0.499	4,858	0.468	0.499
Childless	5,099	0.216	0.412	4,858	0.220	0.414

Table 1.6 shows descriptive statistics for the MZU 71.[21] 52.5% of husbands are blue collar-workers, 21.8% are employees or civil servants up to a middle level. Only 8% of husbands are upper-middle level employees or civil servants and higher.[22] The average score of the husband with regards to social status is 110.57, the average in terms of socio-economic status is 96.20.

When comparing scores of the women's father and her spouse, 27.5% (31.5%) of women have a spouse whose social (socio-economic) status score is lower than their fathers score. Respondents are asked to provide information on their parents for the point in time when they were 15 years of age. We use the father's social/socio-economic status as a proxy for the socio-economic background of a woman's family, since men were typically the main

[21] We cannot provide descriptive statistics on age at marriage or age at first birth with the MZU 71. Year of marriage information was taken over on the household level from the standard Microcensus of 1971. As a consequence, it can only be assigned correctly for married head of households. Also, the MZU 71 does not contain a birth biography.

[22] For married women, the outcome variable contains information on the last job of the husband, for divorced (widowed) women the last job before divorce (death). In case a women is married a second time, the information refers to the current spouse. The MZU 71 does not provide information whether a person has been married previously.

breadwinner at the time. Also, given that the majority of mothers were housewives or unpaid family workers it would not be feasible to assign a social status score to them.[23]

Table 1.6: Summary Statistics Microcensus 1971

		1922 & 1924	
Variable	N	Mean	SD
Year of birth	4,899	1922.98	1.00
Female LFP 1971	4,899	0.479	0.450
Work before marriage	4,402	0.759	0.428
Work after marriage	4,429	0.396	0.489
Social status women before marriage	2,709	85.44	43.54
Husband			
Score social Status (Mayer)	4,386	110.57	53.31
Score social Status (Handl)	4,386	96.20	73.07
Husband has lower score than father of women (Mayer)	3,512	0.275	0.446
Husband has lower score than father of women (Handl)	3,512	0.315	0.465
Husband job	**N**	**Percent**	
Farmer	322	7.3%	
Blue collar worker	2,313	52.5%	
Lower/middle level civil servant/employee	958	21.8%	
Upper/higher level civil servant/employee	354	8.0%	
Self-employed	455	10.3%	

Note: LFP is defined here as employees, self-employed and unpaid family workers.

1.6 Results

In this section, we first present results for social status of the husband and whether women were less likely to marry down. Here, we apply a two sample IV-estimation approach. We then present results on labor market outcomes in later life. Since for this part of the analysis all information is available in one data set, we can apply a standard IV (Wald) estimator.

[23] 64.4% of the mothers are housewives, 13.4% are unpaid family workers and for 11.6% of the cases, information on occupational status of the mother is missing. In the majority of cases when information is missing women did not live with their parents at the age of 15.

1.6.1 Social status of husband - MZU 71

For our two-sample estimation approach, we have to estimate the first stage with year of birth information. The share of women with vocational education increased from 27.0% in 1922 to 34.5% in 1924, hence an increase of 7.5 percentage points (Table 1.7, Row 2). The first stage estimate also reveals whether the instrument has power. With an F-statistic of 32.35, we can have some confidence that our instrument has enough power to detect an effect in case there is any.

Table 1.7: Results social status husband - Census 1970 / MZU71

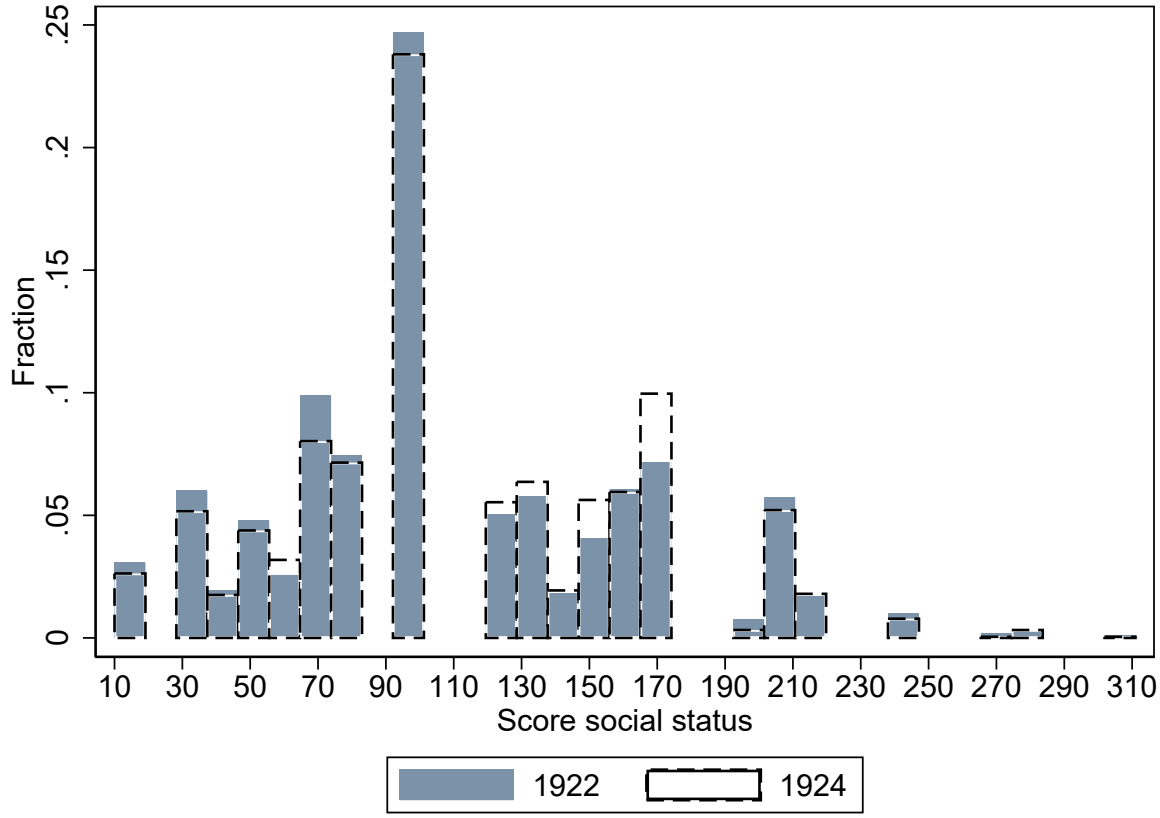
	(1) born 1922	(2) born 1924	(3) Difference [Std. Error]
Social status husband	108.904	112.281	3.376** [1.609]
Vocational Training	0.270	0.345	0.075*** [0.013]
Wald estimate			44.877** [22.790]

For women born in 1922, the average social status score of their husbands is 108.904, the average for cohort 1924 is 112.281. The difference in social status between both birth cohorts and hence the reduced form estimate is 3.376 and is statistically significant at the 5% level (Table 1.7, Row 1).

To investigate which occupational groups drive this increase, Figure 1.2 provides overlaid histograms for social status for the birth cohorts 1922 and 1924. The figure shows that the increase is driven by a higher share of women from birth cohort 1924 having a husband who is a middle level civil servant (score 152) or a middle level employee (score 167). These shares increased from 4.1% in 1922 to 5.6% in 1924, respectively from 7.2% to 10.0%. Correspondingly, the share of women married to a semi-skilled (score 65) or skilled industry worker (score 97) is lower for the birth cohort of 1924. The share of women born in 1924 married to a semi-skilled worker is 8.0% (9.9% for birth cohort of

1922) and the share married to skilled industry workers is 14.3% (15.5% for women born in 1922).

Figure 1.2: Social status of husbands for women born in 1922 versus women born in 1924



We obtain the Wald estimate by dividing the reduced form by the first stage estimate. We find that exposure to vocational training leads to an statistically significant increase in social status of the husband by 44.877 (Table 1.7, Row 3). The mean social status score is 110.57 in our sample. Since none of the occupational categories have exactly a score of 111, we have to take the category closest to the sample mean for illustration of the effect. "Employees in civil service" is the category closest to the sample mean (score of 122). An increase of 45 in terms of social status would correspond to marrying a middle level employee (score 167) instead of an employee in civil service (score 122).

Since the socially desired role for women in the West German society was being a mother and housewife, economic security was an important component of marriage from a women's perspective. Besides, the social status of the husband determined the social

status of the wife and thus her social environment since women often stopped working after they got married. In consequence, women aimed at maintaining social status (or marrying up if possible) relative to the their family background. We compare social status of the husband and the father of the women and look at whether women who completed vocational training were less likely to marry down. The average social status score of fathers for women born in 1922 is 90.18 and amounts to 89.22 for fathers of women born in 1924.[24] This alleviates potential concerns that the fathers of women born in 1922 and 1924 might be systematically different from each other in terms of social status.

We define marrying down as the husband having a lower social status than the father of the women. Row 1 of Table 1.8 displays the reduced form estimate in that regards. 29.1% of women born in 1922 married down compared to 25.8% of the women born in 1924, hence a 3.3 percentage point decrease. We find that women obtaining vocational education are 43.9 percentage points less likely to marry down (Table 1.8, Row 3). While this effect appears huge, it is important to stress that it provides an estimate for compliers, hence a very specific group of people.

Table 1.8: Results probability marrying down - Census 1970 / MZU71

	(1) born 1922	(2) born 1924	(3) Difference [Std. Error]
Social status husband lower than father	0.291	0.258	-0.033** [0.015]
Vocational Training	0.270	0.345	0.075*** [0.013]
Wald estimate			-0.439** [0.214]

[24] The difference is not statistically significant.

1.6.2 Labor market outcomes - Census 1970

We now turn to labor market outcomes in later life. Also here, we begin with assessing the power of the instrument via the first stage (Table 1.9, Row 2). The share of women completing an apprenticeship increases by 7.8 percentage points from cohort 1921/22 to cohort 1923/24 and the F-statistic is 37.81. As stated above we exclude the cohort 1922/23 from the analysis.

With regards to labor market outcomes, we look at whether women are more likely to work in 1970, whether they are more likely to work more than certain amounts of hours per week or whether they earn more. We cannot compare outcomes at the same age, since the information is only available from 1970. There is no evidence that women from cohort 1923/24 are more likely to work or to work more than 15 (25) hours per week than women from cohort 1921/22 as can be seen from Table 1.9. Earning information is unfortunately interval coded and very coarse. In addition, earning information is not available for individuals working in the farming sector.

To illustrate the distribution of individuals across earning categories, we provide overlaid histograms for cohorts 1921/22 and 1923/24 in Figure 1.3. The figure shows that the distribution looks very similar for both cohorts.[25] The category 500DM - 800DM contains the largest number of women for the group of employed women in both cohorts.[26] There is a slight increase in earnings in the range 800-1200 DM from cohort 1921/22 to 1923/24. To investigate this in more detail, we construct a dummy indicating whether a women earned more than 800 Mark. However, this difference is not statistically significant (Table 1.9). Taken together, these results provide no evidence that labor market outcomes were altered by vocational education.[27]

[25] Note, individuals working in the agricultural sector are not included in the figure.

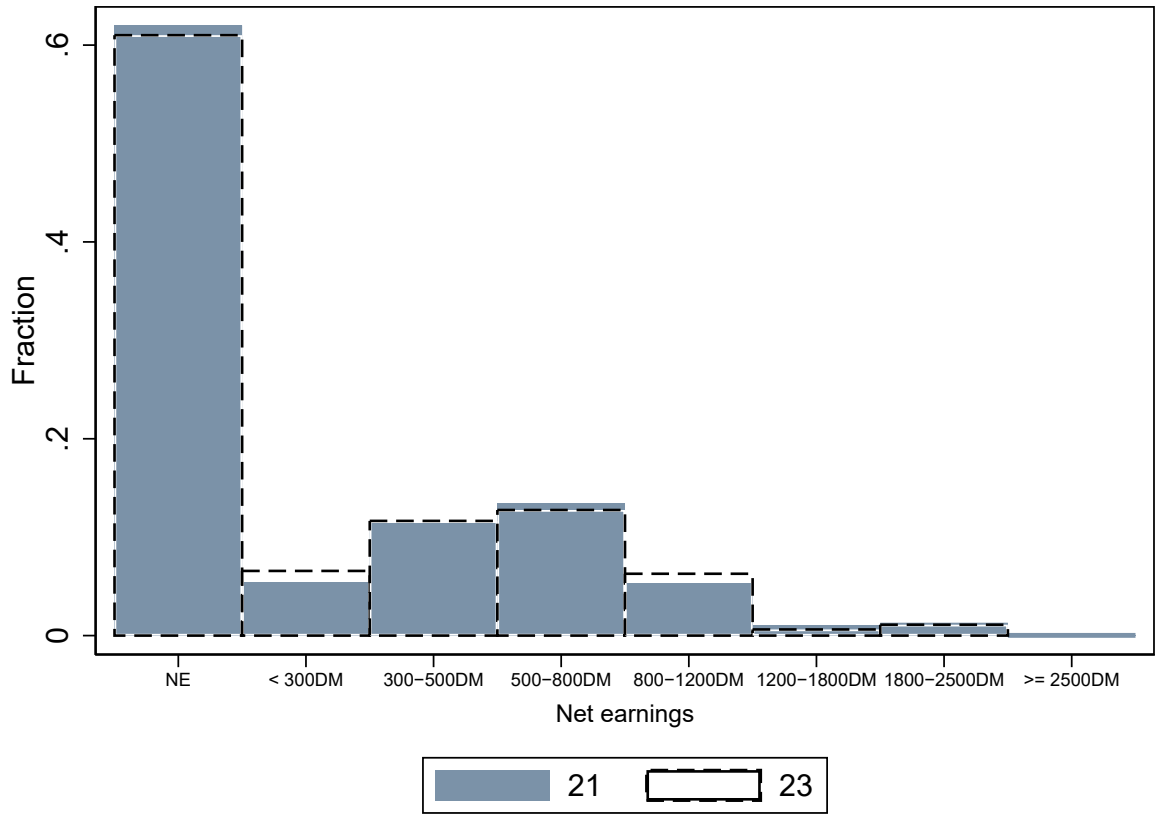
[26] "DM" is the abbreviation for "Deutsche Mark" which is the former German currency.

[27] As a robustness check, we re-estimate the reduced form with $Z_i = 0$ if year of birth is 1922 and $Z_i = 1$ if year of birth is 1924. Also in that case, we do not find evidence that labor market outcomes are altered.

Table 1.9: Labor market outcomes 1970 - Census 1970

	(1) cohort 1921/22	(2) cohort 1923/24	(3) Difference [Std. Error]
Working 1970	0.467	0.478	0.011
			[0.014]
Vocational training	0.254	0.333	0.078***
			[0.013]
Wald estimate			0.135
			[0.179]
Work 15 hours or more	0.449	0.457	0.008
			[0.014]
Wald estimate			0.102
			[0.178]
Work 25 hours or more	0.401	0.399	-0.002
			[0.014]
Wald estimate			-0.022
			[0.175]
Earn more than 800 Mark	0.077	0.080	0.003
			[0.008]
Wald estimate			0.034
			[0.104]

Figure 1.3: Categories of net earnings of cohorts 1921/22 and 1923/24



"NE": No earning information, category includes individuals not working; **"DM"**: Deutsche Mark

1.7 Robustness and channels

Before presenting a potential channel why vocational training impacted partner choice, but left labor market outcomes unchanged, we assess the robustness of our results in the following section.

1.7.1 Robustness

Reduced form estimates of social status score of the husband on fictitious reform

As a first robustness check, we re-estimate the reduced form on placebo reforms to both sides of the actual reform. Since we compare birth cohorts 1922 and 1924 in the main

estimates, we also choose two year differences in this placebo exercise. Our placebo tests are hence 1920 vs. 1922, 1919 vs. 1921, 1918 vs. 1920, 1924 vs. 1926, 1925 vs. 1927 and 1926 vs. 1928. None of the 6 placebo reforms reject the null hypothesis of no effect of the artificial reform on husbands' social status as can be seen in Table 1.10. Since information on parents is only available for birth cohorts 1920 and later, we cannot conduct the placebo exercises 1919 vs. 1921 and 1918 vs. 1920 with regards to probability of marrying down. None of the 4 feasible placebo reforms rejects the null of no effect of the placebo reform on marrying down either (Table 1.10). Taken together, the findings from the placebo exercise make us confident that the significant effects are attributable to a real effect.

Table 1.10: Results reduced form fictitious reforms - MZU71

	(1) Social status husband	(2) Social status husband lower than father
1918 vs. 1920	2.213	
	[1.877]	
1919 vs. 1921	1.912	
	[1.668]	
1920 vs. 1922	-1.981	0.020
	[1.616]	[0.015]
1924 vs. 1926	-0.957	0.014
	[1.641]	[0.015]
1925 vs. 1927	-1.752	-0.002
	[1.687]	[0.015]
1926 vs. 1928	-0.195	-0.002
	[1.681]	[0.015]

Estimation with an alternative score

While we use the social status score developed by Mayer (1977) for the main analysis, we additionally re-estimate our results using the score for socio-economic status proposed by Handl (1977) (Table 1.11 and Table 1.12). As noted above, both scores are highly

correlated. The average socio-economic status of the husband for birth cohort 1922 is 94.762, the average for cohort 1924 is 97.674. While the difference in socio-economic status between both cohorts is similar in size as the difference in social status (for both scores, it is an increase by 3.1%), the reduced form estimate is not statistically significant due to higher standard errors. The reduced form estimate for the probability of marrying down is -4 percentage points and significant at the 5% level, thus corresponds to an 11.64% decreased likelihood of marrying down. This result is similar in size as the estimate with the social status score.

Table 1.11: Results social status husband - Handl score - Census 1970 / MZU71

	(1) born 1922	(2) born 1924	(3) Difference [Std. Error]
Social status husband HSES	94.762	97.674	2.912 [2.206]
Vocational Training	0.270	0.345	0.075*** [0.013]
Wald estimate			38.703 [30.093]

Table 1.12: Results probability marrying down - Handl score - Census 1970 / MZU71

	(1) born 1922	(2) born 1924	(3) Difference [Std. Error]
Social status husband lower than father HSES	0.335	0.296	-0.040** [0.016]
Vocational Training	0.270	0.345	0.075*** [0.013]
Wald estimate			-0.525** [0.228]

Marriage outcomes

Finally, we check whether marriage outcomes are altered. Table 1.13 provides reduced form estimates for ever married, age at first marriage, age difference of partners, number of kids and time span of first child to marriage. While we are able to assess age of first marriage for married, divorced and widowed women, we can only calculate age differences of spouses for existing marriages.

Table 1.13: Reduced form estimates marriage outcomes - Census 1970

	(1) cohort 1921/22	(2) cohort 1923/24	(3) Difference
Not married	0.089	0.089	-0.000 [0.008]
Age at first marriage	25.638	25.693	0.055 [0.151]
Age difference partner (all current marriages)	3.632	3.364	-0.269 [0.179]
Age difference partner (married women, no previous marriage)	3.784	3.424	-0.360** [0.180]
Number of kids	1.838	1.890	0.052 [0.046]
Time difference first child to marriage	1.400	1.478	0.078 [0.102]

We do not find significant differences between both cohorts in terms of ever married, number of children, the time span of birth of first child to marriage, age at first marriage and the age difference of partners (when looking at all current marriages). When we further restrict the analysis to women who have not been married previously (Table 1.13, Row 4), we observe an significant age difference of spouses between both cohorts of -0.36 years, hence around 4 months. Both comparisons with regards to age differences of spouses are unfortunately only coarse approximations of what we would like to measure, namely age differences of spouses for women's first marriages. Nevertheless, since the

age difference is on average 3.5 years (3.6 years when further restricting to no previous marriages), both cohorts marry on average men in cohorts that were fully drafted by the German Wehrmacht (the last birth cohort which was fully drafted is the birth cohort of 1927).

1.7.2 Channels

Prevailing social norms are one potential reason why vocational training did not impact labor market outcomes, but affected partner choice on the other hand. To capture prevalence of social norms on womens' behavior, we look at the time of marriage. Time of marriage is a well-suited point to investigate social norms, since the transition from being unmarried to married is accompanied by drastic shifts in tags about adequate behavior or roles of women. When social norms are strong, one would expect women to behave similarly independent of differences in educational degrees. Specifically, one would expect that even if women have a higher occupational status before marriage, women would leave the labor market equally likely after marriage since they would incur dis-utility from violating the norm.

The MZU 71 provides information on occupational status and sector right before and after marriage. 75.9% of women born in 1922 or 1924 are working before marriage. Since we have information on occupational status before marriage, we can assign social status scores to the women in our sample. We find that being born in 1924 rather than in 1922 translates to an 3.877 increase in terms of social status before marriage (Table 1.14, Row 1). This effect is statistically significant at the 5% level. These results are driven by an decrease in unskilled workers by 3.7 percentage points and an increase in employees by 4.2 percentage points going from birth cohort 1922 to 1924.

Re-scaling the reduced form through division by the first stage, vocational education translates to an increase in social status of 51.524 before marriage (Table 1.14, Row 3). The average social status score is 85.436 for our sample and the estimate represents very roughly the difference in social status between a semi-skilled worker in the service sector and an employee. After marriage, the share of working women drops to 39.6%. Figure 1.4 illustrates the transition in terms of labor force participation around the time of marriage. Exemplarily, *"Yes-Yes"* indicates that a women works before as well as after marriage.

"Yes-No" means that a women works before marriage but is not employed immediately after marriage.

Table 1.14: Social status women before marriage - Census 1970 / MZU71

	(1) born 1922	(2) born 1924	(3) Difference
Social status before marriage	83.575	87.451	3.877** [1.673]
Vocational Training	0.270	0.345	0.075*** [0.013]
Wald estimate			51.524** [24.005]

Figure 1.4: Labor force status around marriage across birth cohorts



"Yes-Yes": Work before and after marriage; **"Yes-No"**: Work before but not after marriage; **"No-No"**: Do not work before as well as after marriage

Despite their better social status right before marriage, women born in 1924 did not work with a higher likelihood after marriage (Table 1.15, Row 1). They were also not more likely to work before and after marriage *("Yes-Yes")* (Table 1.15, Row 2) or less likely to work before marriage and not to work after marriage *("Yes-No")* (Table 1.15, Row 3).

Table 1.15: Labor force participation around the time of marriage - MZU71

	(1) born 1922	(2) born 1924	(3) Difference
Work after marriage	0.402	0.389	-0.014 [0.015]
"Yes - Yes"	0.346	0.337	-0.008 [0.014]
"Yes - No"	0.339	0.343	0.004 [0.014]

Taken together, we interpret these results as an indication that social norms might have played a strong role. Of course, norms might be influential on husbands and wives to different extents. While the view that men with working wives are bad providers might have more impact on men, the attitude that working women are bad wives and uncaring mothers might be more relevant to women (Bernhardt et al., 2018). Unfortunately, we cannot disentangle which might be the driving force for the behavior we observe. We believe that is this an important topic for future research.

1.8 Conclusions

We exploit a German labor market law which strongly increased the supply of available spots for vocational training. Vocational training led women to marry men with a higher social status. Also, they were less likely to marry down when comparing social status of their father and husband. While vocational training did not affect labor market outcomes in later life, it affected partner choice and hence social opportunities through that. We

look at the time of marriage to investigate whether social norms could be one potential channel. We see that while women obtaining vocational education had a job of higher social status before marriage, they were not more likely to stay in the labor market after marriage. We interpret this as an indication that social norms might have played a strong role.

Hence, this paper illustrates that evaluating effects of education might give ex-ante unexpected results due to prevailing norms and ideas. The findings in this paper are consistent with a recent literature that explores the role of social norms and identity in economic exchange.[28] Akerlof and Kranton (2000) introduced social categories, norms and identity into the utility function. They show that this in turn generates new trade-offs people are facing when making decisions. In the West German society, the prevailing norm was the male breadwinner and the women being housewife and mother. Although women with exposure to vocational training had comparably better jobs right before marriage which could have translated into higher consumption possibilities or old-age provisions, being a working wife would violate the appropriate role as mother and housewife. If one assumes that social and economic status is important and considering that a woman's social status was contingent on their husband, it appears that women try to improve their position within prevailing social norms and keeping identity. In our case, exposure to vocational training enabled them to marry a husband with higher social status and hence increased social opportunities through that.

[28] Inter alia, Bernhardt et al. (2018), Bertrand (2011), Bursztyn, Fujiwara, and Pallais (2017) and Steinhauer (2018) study the link on gender identity and female labor force participation.

2 Unbalanced sex ratios in Germany caused by World War II and their effect on fertility: A life cycle perspective

with Iris Kesternich, Bettina Siflinger, and James P. Smith[1]

Abstract

This paper analyzes the effects of permanently unbalanced sex ratios in Germany caused by World War II on fertility outcomes over the life cycle. Using Census records linked with individual biography data, our analysis confirms the commonly found short-term pattern of decreased fertility rates due to a stark imbalance of the sex ratio. Yet, the long-term effects of such an imbalance crucially depend on when in the life cycle fertility is evaluated. We find that female cohorts with low sex ratios have fewer children at younger ages and a larger fraction remains childless. While childlessness remains higher throughout their life cycle, mothers from affected cohorts catch up and even overcompensate at later ages with respect to the number of children. Our preferred reading of this result is that with low sex ratios women select themselves into late motherhood according to their fertility preferences. This interpretation is consistent with the finding that women from affected cohorts with children marry later; they expand their childbearing period; and they accept lower quality matches on the marriage market.

Keywords: Sex ratio; Fertility; Marriage; Life Cycle; World War II

JEL codes: J10; J12; J13

[1]We are grateful to Joshua Angrist, Joseph Zweimueller, Michael Lechner and to participants at seminars at St. Gallen, Passau, Essen, Lausanne, Nuremberg, Mainz and at the Labor Workshop in Zams. We would like to thank Ferdinand Kramer for providing expertise on the historical background and employees from various statistical offices in Germany for their assistance in collecting the historical data. This paper uses data from a series of German Censuses and the SOEP. The research was funded by a grant from the National Institute on Aging (R37-AG025529) to James Smith. Iris Kesternich is grateful to funding from KU Leuven research funds. Bettina Siflinger acknowledges financial support from the German Science Foundation (DFG) through CRC 884. Carina Steckenleiter is grateful for funding from the Swiss National Fund for Scientific Research (SNF) (SNF 100018E-161598).

2.1 Introduction

In past decades, sex ratios were identified as important determinants of demographic and economic outcomes as diverse as gender roles and earnings inequality between males and females (Acemoglu, Autor, & Lyle, 2004). A sex ratio is defined as number of men per women in the population with a normal equilibrium value close to one. Sex ratio imbalances were shown to have significant consequences for labor and marriage markets (see Abramitzky, Delavande, & Vasconcelos, 2011; Angrist, 2002; Bethmann & Kvasnicka, 2013). Unbalanced sex ratios change the bargaining position of men and women in the marriage market and shift intra-household resources and family structures. While nowadays sex ratios are balanced in Western societies, the availability of new technologies of gender control in utero creates increasing imbalances in the ratio of men and women in countries where boys are favored over girls. In particular, India and China face the challenge of counteracting an increased skewness in sex ratios and its demographic and economic consequences (Anukriti, 2018; Jayachandran, 2017).

In this paper, we investigate the long-term effects of highly unbalanced sex ratios in Germany caused by World War II (WWII) on fertility of West German women over the life cycle. While existing papers tend to study effects at one point in time or in the short run, we analyze effects over the life cycle. We document the importance of the life cycle perspective to show that conclusions on the impact of an unbalanced sex ratio crucially depend on when outcomes are measured. We also test whether commonly documented short-term fertility patterns persist over the life cycle, or whether there are offsetting mechanisms at work. Some facts taken for granted – e.g. that lower sex ratios lead to lower marriage probabilities and lower number of children – might not hold up when one follows women over the life cycle. One potential reason is that women have different margins of adjustment to achieve desired fertility outcomes: postponing marriage and fertility, extending childbearing to later ages, or accepting lower quality of matches on the marriage market.

An important innovation of our study is that we decompose total fertility into its extensive and intensive margin. The extensive margin refers to the share of women remaining childless, and the intensive margin denotes the total number of children per mother. The

relationship of those two margins of fertility is crucial to understand the impact of low sex ratios on total fertility, as they could work in the same direction, or they could offset each other (Aaronson, Lange, & Mazumder, 2014; Baudin, De la Croix, & Gobbi, 2015). While a low sex ratio may imply that more women end up childless, the consequences on the number of children per mother is less clear. One potential mechanism is selection along fertility preferences: Under a scarcity of men, women with a strong preference for children sort into (late) motherhood, while women with low fertility preferences decide to remain childless.

To identify the impact of the long-lasting imbalance in sex ratios on fertility over the life cycle, we exploit a stark drop in sex ratios caused by the large share of men in drafted cohorts dying in WWII. While male cohorts from 1910-1927 were fully drafted and faced a mortality rate up to 38% per cohort, later cohorts remained unaffected. Our identification strategy exploits differences in sex ratio across female birth cohorts 1919-1944 within the same region and links them to life cycle fertility. Due to the special situation in post-WWII Germany, factors such as internal migration, the inflow of refugees or the selective loss of population and physical capital do not invalidate our identification strategy. Thus, the variation in sex ratios is outside of individual control. Abramitzky et al. (2011) uses changes in sex ratios due to World War I to analyze changes in mating patterns. Angrist (2002) uses changes in immigrant sex ratios in the US to identify effects on labor and marriage markets. Our source of variation is closest to Brainerd (2017) who uses a sharp drop in sex ratios to identify effects of WWII in the Soviet Union on marriage and fertility.

For our empirical analysis, we use sex ratios from German Census records that vary by regions and by birth cohorts. The main sex ratio measure stems from the first Census after WWII in 1946. We moreover use sex ratios from 1939 to show that the drop in sex ratios were indeed caused by WWII, and sex ratios from 1950, 1961 and 1970 to illustrate that the imbalance in sex ratios is permanent for affected cohorts. The Census records on sex ratios are merged to the German Socio-Economic Panel Study (SOEP), which provides detailed biographic information on marriage and fertility over the life cycle. In addition to survey data, we make use of a 1% sample of the West German Census population in 1970 and merge it with records on sex ratios. Individuals in this sample received an additional survey module to collect comprehensive information on their socio-economic

and demographic characteristics. The data allow us to analyze a number of potential margins women may use to adjust to the situation of low sex ratios.

In the empirical analysis, we first investigate the effects of the permanent drop in sex ratios on women's total fertility. We find that women facing low sex ratios are significantly less likely to have children in their 20s, which is in line with the short-term effects found in the literature (for instance Abramitzky et al., 2011; Bethmann & Kvasnicka, 2013). At later ages, total fertility does not significantly differ between cohorts affected by the scarcity of men and unaffected cohorts. This suggest that women with low sex ratios catch up in total fertility in the longer run. The found fertility pattern is accompanied by a short-run delay in first marriage. Interestingly, we do not find equivalent fertility effects for East Germany that we analyzed separately due to the different political and institutional situation after WWII.

We next separately analyze the intensive and the extensive margin of fertility. While the impact of male scarcity on childlessness is large and permanent, a lower number of children at younger ages is followed by a catching up and even overcompensation, leading to a significantly higher number of children at completed fertility. Using the 1% sample of the West German 1970 Census population, we explore the potential margins of adjustment women facing low sex ratios may use to achieve the desired fertility outcome. We find that a low sex ratio postpones marriage and childbearing age, and expands the childbearing period. Moreover, male scarcity leads to worse matching in the marriage market (as measured by larger spousal age gaps and worse educational assortative mating). We do not find significant effects on out-of-wedlock childbearing.

Our findings lend themselves to the following interpretation. The scarcity of men induces a selection of women into marriage and motherhood along their fertility preferences. Women with strong fertility preferences do marry and have many children, albeit at later ages and at the cost of a potentially bad spousal match. A lower spousal quality may increase the price for child quality and increase fertility along the intensive margin due to a substitution of quality into quantity of children (Aaronson et al., 2014; Becker, 1974; Guinnane, 2011). For women with weak fertility preferences the costs associated with having a worse matching quality are too high to compensate the benefits of marriage and motherhood. They decide to remain childless. The result of this selection is the observed

relationship between the intensive and the extensive margin of fertility. Importantly we observe this relationship over women's life cycle up to completed fertility that is required to understand how fertility reacts to large and permanent population shocks.

We contribute to the literature on unbalanced sex ratios in several ways. Compared to previous studies (Abramitzky et al., 2011; Bethmann & Kvasnicka, 2013), we document that an unbalanced sex ratio not only has immediate consequences but also affects fertility in the long run. Importantly, the magnitude of these effects crucially depend on when fertility is measured in the life cycle. By analyzing the intensive and the extensive margin of fertility separately up to completed fertility, we gain a deeper understanding on how fertility reacts to a shock in sex ratios in the long run. Taking account of these margins of fertility also allows us to investigate other possible mechanisms of sex ratio effects. In constrast to other studies in the literature, our fertility patterns do not seem to be driven by a significant increase in out of wedlock births (Bethmann & Kvasnicka, 2013; Brainerd, 2017).[2] Instead, they are due to a postponed and prolonged period of fertility and a worse match quality. Finally, while other studies were using transitory changes in sex ratios for the analysis of long-term effects (Acemoglu et al., 2004; Goldin & Olivetti, 2013), we document the consequences of a permanent reduction in the number of men per woman.

The remainder of the paper is organized as follows. The next section provides a brief summary of the existing literature. Section 2.3 presents the Census data and describes the historical context as well as the variation in the sex ratios. Section 2.4 describes the individual level data and provides descriptive statistics of the data we will use in this analysis. Section 2.5 contains our models of the effects of these changing sex ratios on demographic outcomes. Results are presented and discussed in Section 2.6. The final Section 2.7 highlights our main conclusions.

2.2 Related literature

According to the marriage model (Becker, 1973), a decrease in sex ratio improves male bargaining position increasing male marriage rates and economic resources. Several au-

[2] A potential explanation is that non-marital births were highly stigmatized and non-marital children were under guardianship of the court if the mother could not prove marriage at birth.

thors studied short-term consequences of reductions of men per women. Abramitzky et al. (2011) document effects of changes in sex ratios in France caused by male mortality in WWI. They show that regions with a larger sex ratio decrease, men were more and women were less likely to marry three years after end of WWI. Divorce rates decreased and out-of-wedlock births also increased.

Lafortune (2013) analyzes sex ratios in second generation American immigrants and shows that greater scarcity of potential partners leads individuals to invest more in attributes considered attractive by potential partners. Using county-level Census data for the German state of Bavaria for 1939 and 1946, Bethmann and Kvasnicka (2013) document that reductions in number of men per women due to WWII increases the probability of non-marital fertility. This effect is stronger for counties with a lower expected return of soldiers. Brainerd (2017) uses both sex ratios and outcomes measured in the first Russian post-war Census of 1959 to document that dramatic drops in sex ratios for several cohorts and regions in Russia lead to lower rates of marriage and fertility as well as higher rates of out-of-wedlock births and divorces for women.

There is a literature documenting effects on female labor supply. Angrist (2002) showed that lower male sex ratios among second generation immigrants in the US lead to higher labor market participation and lower female marriage rates. Acemoglu et al. (2004) find that in US regions where more men were serving in WWII, more women entered the labor market, leading to a lowering in female wages and increased earnings inequality. There is an unresolved discussion on whether women who worked due to WWII mobilization did so permanently or whether they became housewives when soldiers returned from WWII. At least for some groups (those entering white-collar jobs), the change was permanent (Goldin & Olivetti, 2013). Effects of increasing labor supply even carry over to the next generation by making sons of working mothers more likely to have working wives themselves (Fernández, Fogli, & Olivetti, 2004). The latter two studies considered long-run effects of unbalanced sex ratios. These papers do not document life-cycle effects on individual women and cannot make inferences on age-specific patterns of unbalanced sex ratios. In the US labor market, men absent during the war were replaced by women. In Germany, this replacement was mainly done by prisoners of war and forced laborers, the labor force participation of German women increased only slightly over the

course of the war (Kaldor, 1945).

The recent literature deals with different processes driving the extensive margin (whether someone becomes a mother) and intensive margin (fertility of mothers) of fertility. Baudin et al. (2015) argue that two forces driving these margins are poverty (poor women not having access to technologies overcoming childlessness) and opportunity costs (higher for highly educated women). Aaronson et al. (2014) show that allowing for an extensive margin of fertility in the Beckerian quantity-quality model generates new insights for fertility transitions. Changes in the price of quality of children can influence intensive and extensive margin differentially.

Lastly, this paper relates to recent studies on long-term effects of war. This literature stresses that there are negative long-term effects on health and labor market outcomes (Akbulut-Yuksel, 2014; Kesternich, Siflinger, Smith, & Winter, 2014), and on preferences of the surviving population (Bauer et al., 2016). Regarding WWII, the mostly studied channels of how wars affected well-being of survivors has been the experience of hunger (Jürges, 2013; Kesternich, Siflinger, Smith, & Winter, 2015; Lumey, Stein, & Susser, 2011; Neelsen & Stratmann, 2011).

2.3 Sex ratios and historical context

2.3.1 Sex ratios constructed from German Census data

We collected German Census records from the first Census after WWII in October 1946. The Census was conducted by the four occupational powers (USA, UK, France and Soviet Union) for remaining territories of Germany. By constructing sex ratios at German state level, we obtain sex ratios that vary across birth cohorts and state of residence. Sex ratios are heterogeneous across different German states. To assess pre-war sex ratios and to compare them with post-war sex ratios, we collected sex ratios from the Census of 1939.[3] Finally, we obtained data on the number of expellees in 1946 by year of birth for the states of Bavaria, North-Rhine Westphalia (NRW) and states in the Soviet zone to investigate

[3] A challenge in analyzing the data from 1939 is the different regional structure of Germany before and after WWII. Since we analyze Germany in todays' borders, "provinces" of 1939 are matched to the post-unification states of Germany. A detailed description of regional differences before and after WWII and the matching procedures are contained in Appendix 2.B.

their influence on gender ratios.[4]

To analyze long-term patterns in sex ratios, we collected sex ratios from Census years 1950, 1961 and 1970 in West Germany. From 1949 onwards, the three western occupation zones formed the Federal Republic of Germany ("West Germany") and the Soviet zone became the German Democratic Republic ("East Germany"). For the relevant period, most population data by birth year and state level have not been digitally available. We obtained data on sex ratios by gender, birth cohort and German states from State Statistical Offices, the Federal Statistical Office of Germany and from various printed publications. Detailed references for all Census years and states are contained in Appendix 2.B. To our knowledge, we are the first assembling a data set with sex ratios by birth cohorts on the state level for Germany for these five Census years.[5]

2.3.2 Variation in sex ratios

Sex ratios are typically defined as the number of men divided by number of women stratified by year of birth and region of residence. Figure 2.1 displays two subfigures for sex ratios in Census years 1939 and 1946. Figure 2.1.A shows that in 1946 sex ratios strikingly vary over birth cohorts as well as across region of residence.[6] While sex ratios for those born after 1927 (about 18 years old at the end of WWII) are close to unity, there was a dramatic fall in sex ratios over all regions for those born before 1928. This pattern can be explained by drafting pattern of German armed forces: Over the course of the war, 17.3 million men were drafted in total.

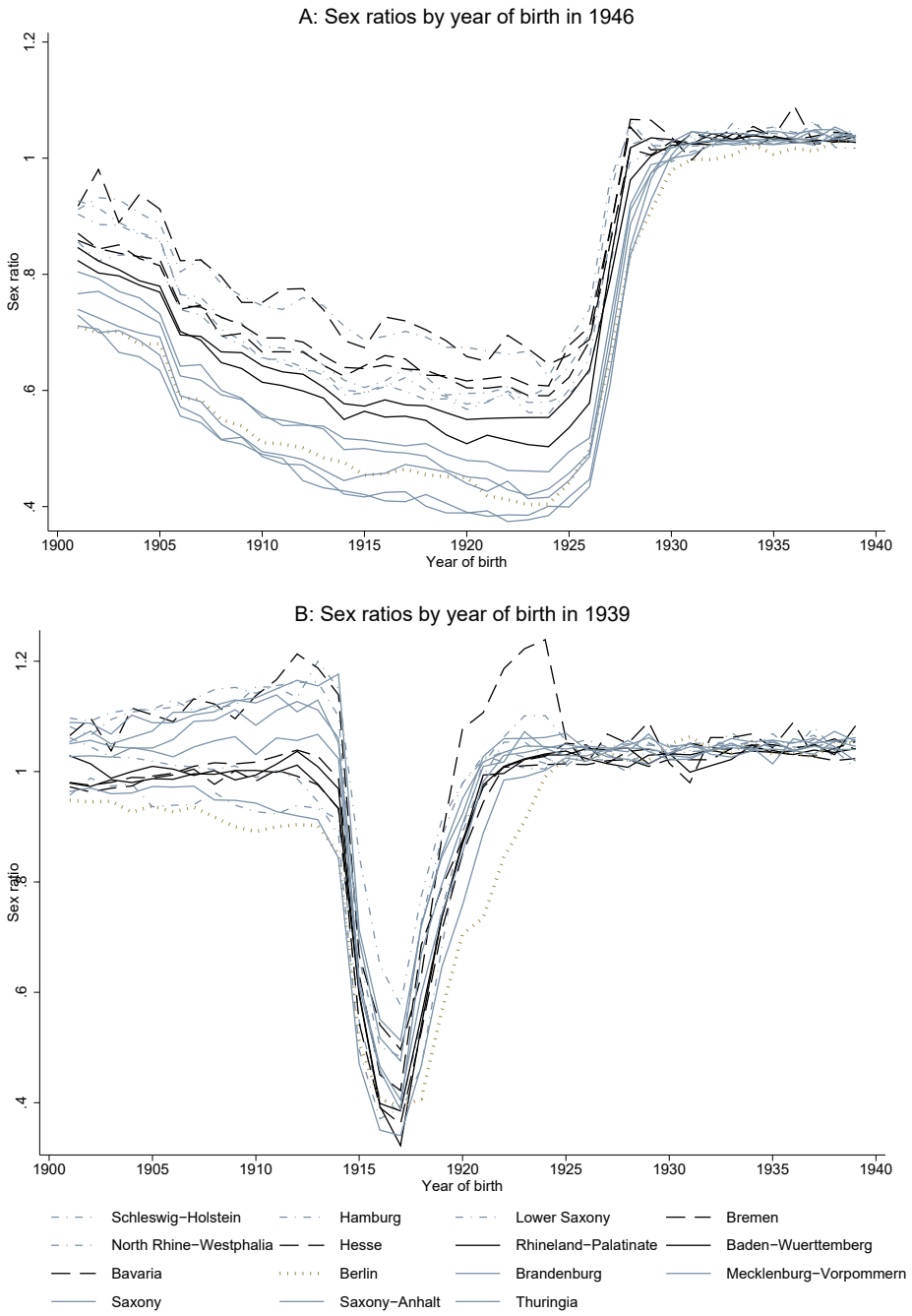
While birth cohorts from 1910 to 1927 were fully drafted, those born in 1928 and 1929 were only drafted occasionally (Kroener, Müller, & Umbreit, 1988). Figure 2.2 shows the fraction of German men in each birth cohort who died in the war. While in total about 5.3 million German soldiers died during the war (Overmans, 2004), about 34%-38% of men of birth cohorts 1916–1924 died in war. This fraction declined dramatically after

[4]The term expellees refers to migrants from the former Eastern territories of the German Reich or Ethnic Germans who lived as minorities in countries such as Czechoslovakia. Census results of 1946 include expellees.

[5]Note that since we want to make use of the stark variation in sex ratios over birth cohorts, we cannot use data on smaller geographical units than German States, since the Census data is either disaggregated by regions within a State or by cohorts. In a related study, Grant, Kesternich, Steckenleiter, and Winter (2018) use sex ratios at the district level ("Landkreise").

[6]See also discussion in Grant et al. (2018).

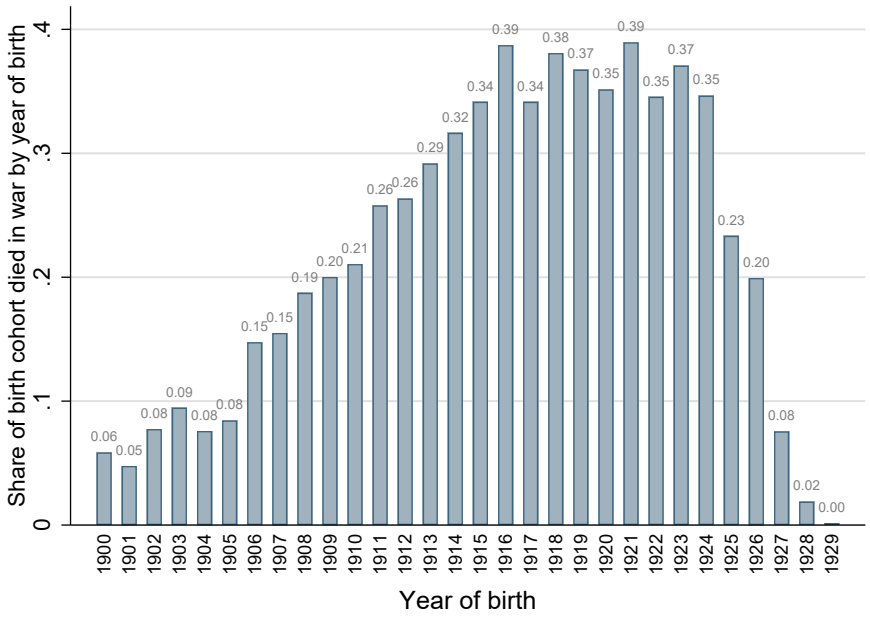
Figure 2.1: Sex ratios by year of birth in 1939 and in 1946



Note: Soviet zone states refer to grey solid lines; US zone states refer to black dashed lines; French zone states refer to black solid lines; UK zone states refer to grey dashed-dotted lines; Berlin is the dotted line. UK zone: Census data record all individuals attending this zone during data collection. Other zones: Census data collected for resident population. Population in all zones includes camp inmates, but excludes camp inmates of displaced person camps. The content of this figure is also displayed in similar representation in Grant et al. (2018). *Data source:* References for Census data are presented in Appendix 2.B.

1925 and more gradually for birth cohorts before 1914.

Figure 2.2: Share of male birth cohort that died in war by year of birth



Data source: Kroener et al. (1988), p.986

Figure 2.1.A shows considerable differences in sex ratios in 1946 across German states. States within the Soviet zone (solid lines in blue) and Berlin (which was divided among the allied forces) were particularly affected by lack of men in 1946. States within the British zone (short dashes and dots alternating in blue) as well as the US zone (long dashes in black) display much higher sex ratios compared to Soviet and French zones (solid lines in black).[7] Geographical clusters in sex ratios outlined in Figure 2.1.A were largely a result of different conditions at the eastern and western frontlines at the end of the war.[8] While the German Army was pushed back on all fronts as Allied troops advanced, large numbers of German soldiers on the Eastern front tried to flee to the West to escape imprisonment by the Soviets (Henke, 1996). By June 1945, about 5 million had surrendered to the US army, and it is estimated that about 1 million of them fled from the Soviets (Bischof & Ambrose, 1992). Battles in the last days of war were much fiercer in the Soviet zone. To hold back the Soviets, available men from the local population including

[7]The state Baden-Wuerttemberg in its today's borders was formed out of three smaller states in 1952. Two of the former states were part of the French zone, one state was part of the US zone.

[8]We also include the states that became part of East Germany in the figures displaying the sex ratios in 1939 and 1946 to give a better illustration of the historical context.

very young boys and old men had to fight as a last resort without proper training and equipment. This strategy led to very high losses among the male population (Steinberg, 1991). Importantly, the influx of expellees did not considerably alter sex ratios on the state level. Figure 2.3 compares sex ratios of the total population with sex ratios of the native population (without expellees) for Bavaria, NRW and Soviet zone states.

1939 sex ratios are displayed in Figure 2.1.B. The patterns differ from that of the sex ratios in 1946 in two important ways. First, there is a much smaller regional variation across German states. Second, Figure 2.1.B exhibits a sharp but temporary drop in the sex ratios in 1939 for birth cohorts 1915-1920. These birth cohorts were drafted for compulsory military service or labor service in 1939 and hence were not included in the Census (Statistisches Reichsamt, 1940). Patterns in Figure 2.1.B alleviate concerns about selective drafting patterns across German states.

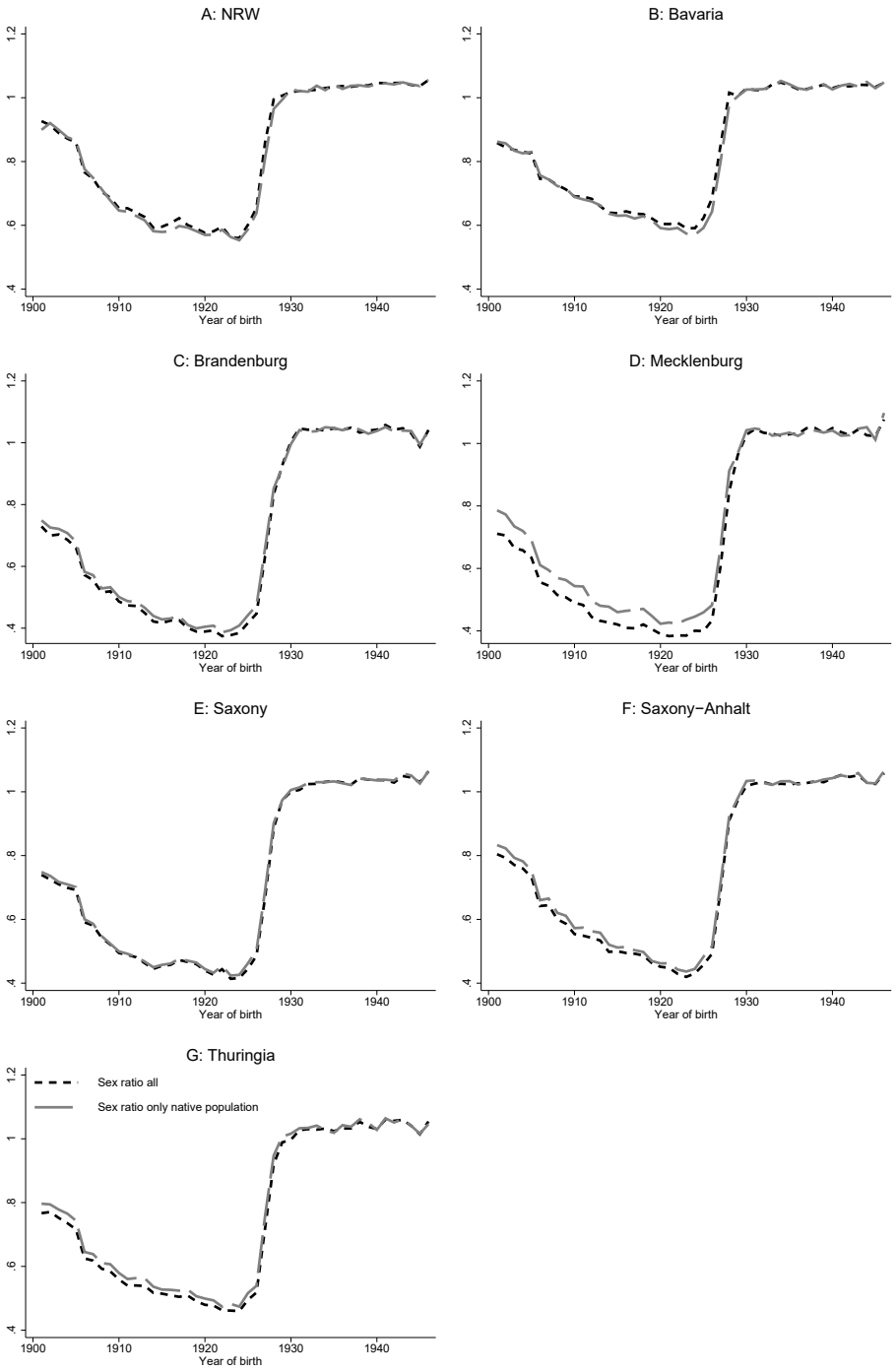
For reasons outlined above, we assume sex ratios changed permanently due to WWII. Figure 2.4 illustrates sex ratios for later Census years. While a decrease in the sex ratio imbalance appears in 1950 due to the return of male prisoners of war (Figure 2.4.A),[9] the sex ratio does not return to its original equilibrium. Figures 2.4.B and 2.4.C show that the sex ratios are permanently lower and had not recovered from the shock of WWII up to 25 years later. In our main analysis, we will use the sex ratios from 1946 since part of the shifts in later years may be endogenous. We present results using sex ratios from 1950 in Section 2.6.3. Persistence of low level of sex ratios due to WWII is corroborated by Figure 2.5, which plots sex ratios by Census years for affected and non-affected cohorts. It shows that individuals born after 1927 face a balanced sex ratio over the whole life course.[10] Sex ratios for individuals born in 1927 and earlier show a dramatic drop between 1939 and 1946. They slightly improve until 1950, and remain highly imbalanced in later years.[11]

[9] Around 1.2 Million prisoners of war returned to West Germany between the Census of 1946 and 1950 (Statistisches Bundesamt, 1950).

[10] For instance, the birth cohort of 1932 permanently faced balanced sex ratios of about 1.

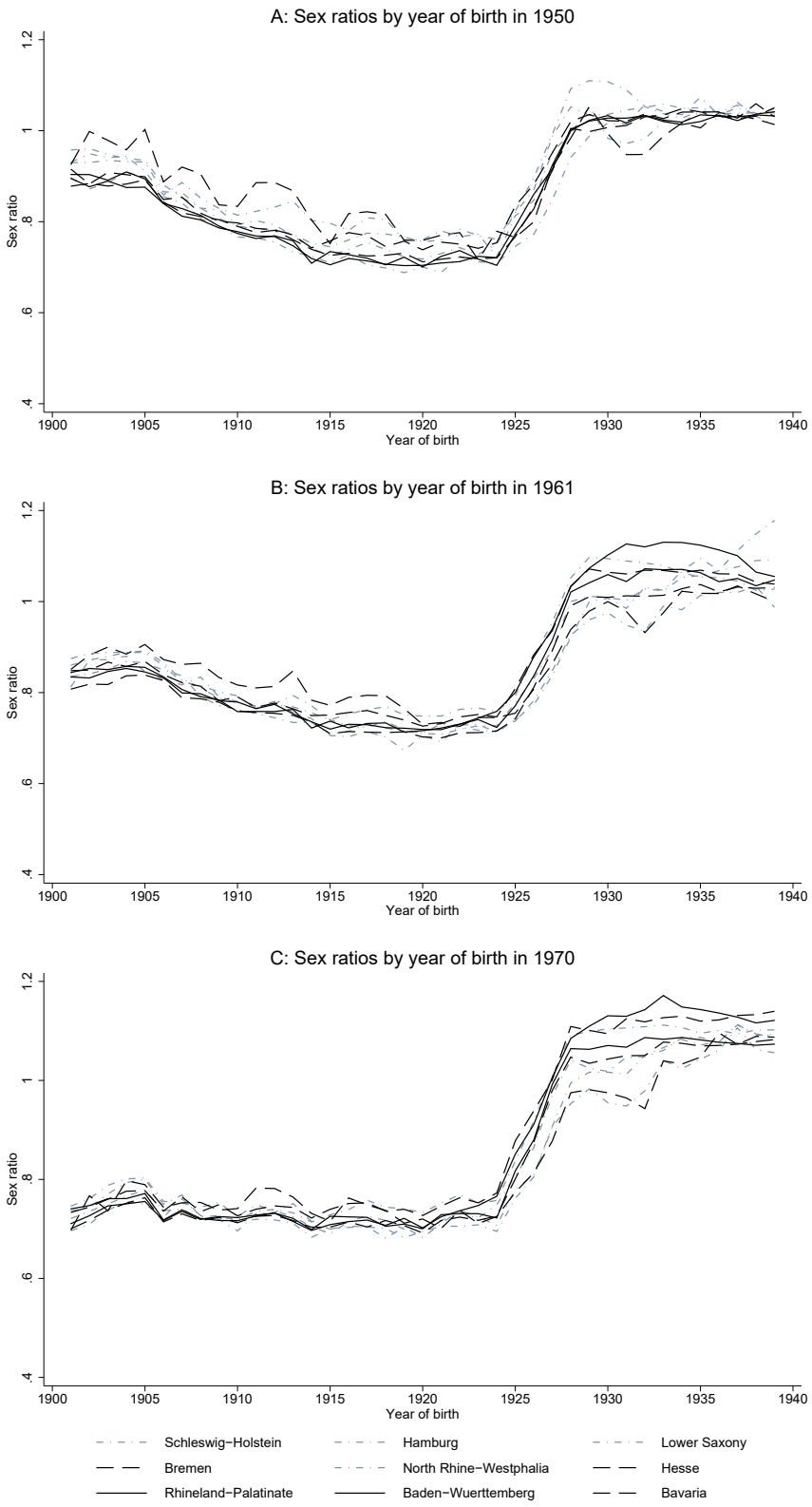
[11] In Figure 2.5, the most affected cohort is individuals born in 1924 whose sex ratio falls to 0.6 and never increases to more than 0.7 over the life cycle.

Figure 2.3: Sex ratios of overall population and native population on state level in 1946



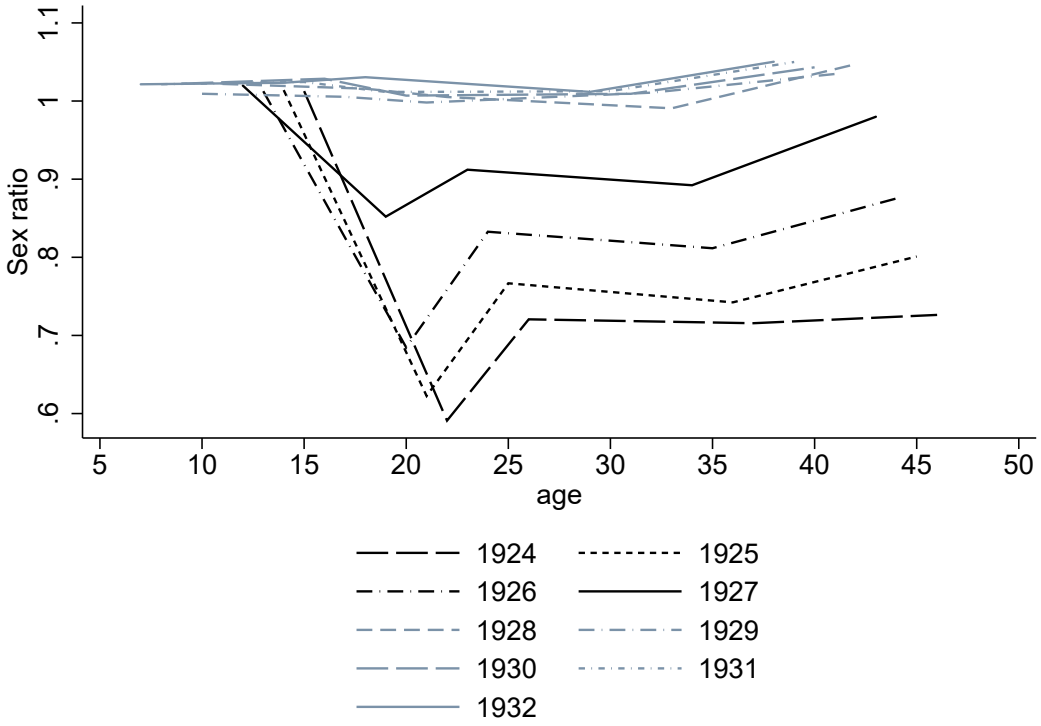
Note: References for Census data are presented in Appendix 2.B. This figure is also contained in Grant et al. (2018).

Figure 2.4: Sex ratios in 1950, 1961 and 1970, West Germany



Note: References for Census data are presented in Appendix 2.B.

Figure 2.5: Persistency in sex ratios over the life cycle for birth cohorts 1924-1932, state of Bavaria



Data source: Census data 1939, 1946, 1950, 1961 and 1970. *Note:* References for Census data are presented in Appendix 2.B.

2.4 Individual level data

2.4.1 The German Socio-Economic Panel

In the empirical analysis, we combine Census data on sex ratios on the German state level and rich individual level data from the German Socio-Economic Panel (SOEP). SOEP is a longitudinal panel survey modelled after the PSID that surveys a representative sample of the German population (Wagner, Frick, & Schupp, 2007). Since 1984, SOEP interviews about 20,000 individuals from more than 11,000 German households, collecting comprehensive respondent information on labor force participation, economic status, health, life history and household composition. For our demographic outcome measures, we use a spell-based biography module which is first collected when a respondent enters the sample and is updated in each wave. This module provides comprehensive annual

information on marital and fertility history of individuals over the life course. Respondents and non-respondent household members report the onset and duration of marriage, divorce, widowhood, and childbirth from age 15 to the interview date. Using this information, we construct marriage and fertility outcomes. Since we are interested in impacts of sex ratios around WWII, we restrict our sample to women born between 1919 and 1944. We exclude individuals from older birth cohorts because they were also affected by WWI (1914-1918) providing us with 6,878 women. We restrict data to respondents who were either born within today's borders of the Federal Republic of Germany (West Germany) or immigrated to this area before 1949. We exclude individuals from West-Berlin and the Saarland because of their special status in the years after WWII and up to the reunification in 1989.[12] This leaves us with a sample of 4,505 women.

To link 1946 Census records with SOEP data, two identifiers are required. We use women's birth year to link sex ratios from birth cohorts. The second identifier is state of residence where women were born or grew up. While this is not available from a direct question in SOEP, we use combined information from two separate questions to proxy women's residence during childhood. First, respondents report federal state of last school attendance available for 1,313 women. Second, respondents are asked if they still live at the place where they were raised or if they moved back to the place of childhood. This information is available for 2,039 women. We combine information from these questions and obtain a proxy for state of residence during childhood for 2,681 women born between 1919 and 1944.[13] Table 2.1 represents a comparison of our analytic sample with total sample of West German women in relevant age group. Means and standard deviations for analytic variables are very similar in both samples. Table 2.1 provides an overview for main descriptive statistics. On average, women in our sample are born in 1935; 93% of these women get married, and 87% have children. They are 24.5 years old when they marry and give birth to their first child about 11 months later, at age 25.4. Unmarried women are almost three years younger when they give birth to their first child. A valid issue with our definition of the analytic sample is that some women may have moved immediately

[12]Since the Saarland was not part of any occupation zone in October 1946, no Census results for 1946 are available.

[13]If respondents report valid answers to both questions, we use information on residence of last school attendance. Kesternich, Smith, Winter, and Hörl (in press) compare this with reliable state of birth obtained from wave 2012 and find it is consistent for about 87% of SOEP respondents born between 1929 and 1955.

Table 2.1: Comparison of analytic sample to sample of West German women: outcomes and control variables

	Analytic sample			Sample West German women		
	N	mean	SD	N	mean	SD
Marriage						
Ever married	2,681	0.93	0.25	4,505	0.93	0.26
Age first marriage	2,500	24.49	6.05	4,170	25.76	8.34
Fertility						
Extensive margin: any children	2,681	0.87	0.33	4,505	0.83	0.37
Intensive margin: number children	2,343	2.33	1.22	3,748	2.31	1.23
Total fertility	2,681	2.04	1.38	4,505	1.92	1.41
Age at first birth if married	2,086	25.35	4.35	3,291	25.53	4.46
Age at first birth if unmarried	182	22.41	4.39	328	22.54	4.3
Controls						
Year of birth	2,681	1935.41	6.78	4,505	1933.91	7.18
Father has upper secondary educ.	2,681	0.08	0.28	4,505	0.08	0.27
Mother has upper secondary educ.	2,681	0.02	0.16	4,505	0.02	0.15
Years of schooling	2,426	10.88	2.19	4,121	10.82	2.14
State of residence childhood						
Schleswig-Holstein	2,681	0.05	0.22			
Hamburg	2,681	0.03	0.17			
Lower Saxony	2,681	0.13	0.33			
Bremen	2,681	0.01	0.11			
North Rhine-Westphalia	2,681	0.3	0.46			
Hesse	2,681	0.09	0.28			
Rhineland-Palatine	2,681	0.05	0.23			
Baden-Wuerttemberg	2,681	0.14	0.35			
Bavaria	2,681	0.2	0.4			

after WWII. If this were the case, our results may be explained by (endogenous) selection rather than by the variation in sex ratios. However, we can exclude such an issue since moving was heavily restricted after the war. Allied forces enforced a law which forbade moving within and across zones until summer 1947. Afterwards, moving was restricted by the Allied Control Council Law No 18 (to obtain a moving permission housing space needed to be proved) as well as by moving bans for many cities (Hottes & Teubert, 1977). In-migration towards these cities was banned until the end of the 1940s as German cities were facing a severe housing shortage in consequence of aerial bombing during the war (see e.g. Kift, 2008).

2.4.2 1 Percent sample of West German population

While SOEP provides us with individual life cycle information, it is lacking information on husbands. To analyze potential mechanisms linking imbalances in sex ratios and fertility, we exploit individual level data from a 1% sample of the West-German population in 1970.[14]

This sample received an extra module with detailed questions on individual's marriage, fertility, labor and other SES attributes, and includes information on spouses. This extra module has two drawbacks compared to SOEP. First, while SOEP tracked women over their whole life course and until completed fertility, the Census records women's status quo in 1970. Second, most women have not reached the age of completed fertility when surveyed. Thus, while we cannot analyze total fertility or life cycle effects using this sample, we can use the data to conduct an analysis of the potential mechanism linking imbalances in sex ratios and fertility patterns.

The original 1% sample consists of 623,131 individuals born between 1870 and 1970. To obtain a sample comparable to that of SOEP, we restrict data to female birth cohorts 1919-1944, living in West Germany before WWII. We exclude Non-German residents, expellees, refugees from territories of former German Reich and only keep women who are household heads or spouses of household heads. We link the resulting sample with 1946 Census records and obtain an analytic sample of 66,106 women. We construct a second sample containing education and marriage information on wives and husbands, resulting in a sample size of 58,954 couples. Using year of birth and education, we compute differences in age and education between wife and husband.

Table 2.2 provides statistics of variables for individual and couple samples. 95% of women are married in 1970 and their first marriage takes place at about age 24 - about 84% of women have children and average number of children is 2.21 along the intensive margin of fertility. Average total fertility is 1.84 children per woman. Age at birth of first child varies by marital status. While married women give birth for first time at about age 25, age for out of wedlock births is about three years earlier, around age 22. By comparing these numbers with descriptive statistics for SOEP in Table 2.1, we find similar values for

[14]The data comprise a 10% sample of residents in West Germany from the German Census in 1970. For a detailed description of the data see (Schimpl-Neimanns & Frenzel, 1995).

age at first marriage and first birth. Fertility numbers are smaller than in SOEP samples, because for part of the sample fertility has not yet been completed.

Table 2.2: Descriptive statistics 1% percent sample German Census 1970

	N	Mean	SD
Individual sample			
Marriage			
Ever married	66,106	0.95	0.2
Age first marriage	59,132	23.79	4.08
Fertility			
Extensive margin: have any children	66,106	0.84	0.37
Intensive margin: number children	55,186	2.21	1.26
Total fertility	66,106	1.84	1.41
Age at first birth if married	50,115	25.16	4.14
Age at first birth if unmarried	5,070	21.96	3.51
Out of wedlock births	55,186	0.082	0.29
Year of birth	66,106	1931.52	7.5
Couple sample			
Wife: year of birth	58,954	1931.84	9
Husband: year of birth	58,954	1929.12	7.39
Wife: years education	58,948	8.85	3.02
Husband: years education	58,692	9.81	2.04
Husband-wife: Difference in age	58,954	3.3	4.9
Husband-wife: Difference in years education	58,687	0.94	2.71

2.5 Empirical strategy

WWII caused a dramatic and sudden drop in the availability of men, resulting in a significant imbalance in sex ratios. Independently of which age our respondents had at the time of the Census, sex ratios measure the drop in number of men that women faced. If someone was too young in 1946 to be affected by drafting of men in WWII, their sex ratio will be close to unity. This individual is unaffected by over–proportional dying of men during WWII. Such a person will face balanced sex ratios all through her life as there were no further big shocks to sex ratios in Germany later. By exploiting this source

of exogenous variation, we estimate empirical models of sex ratios and marriage linking them to fertility and marriage behavior.

Note that sex ratios vary by cohort and region (state). To control for time and cohort-specific trends, we use region and 5–year cohort dummies. The 5–year birth cohort intervals are 1919-1924, 1925-1929, 1930-1934, 1935-1939, and 1940-1944. We cannot include year of birth fixed effects, since our sample in the SOEP is too small to allow for both cohort and region fixed effects. However, absent of the effect of sex ratios, we would not expect any stark changes in fertility within these 5-year cohorts. We will provide several pieces of evidence to confirm this assumption. In particular we will additionally control for whether a specific fertility period falls within WWII; take account of region-specific trends by interacting calendar time and region fixed effects; and add measures of GDP and economic uncertainty to our main specification. Given the stark drop of sex ratios for the cohorts born before 1928, we moreover split the birth cohorts at this cutoff date and used this in the regressions instead. As we will show in the robustness section (Section 2.6.3), our main results also go through in this setting. However, for reasons specified in the next paragraph, we do not believe that such a stark cut-off is the right way to think about our variation.

While sex ratios are typically defined as men/women from same region and birth cohort, individuals can substitute towards partners a few years older or younger. On average, men are about 3.3 years older than women (see Table 2.2). To allow for this type of substitution in our empirical analysis, we define the sex ratio by choosing a window including birth cohorts from up to three years older to two years younger than the respective birth cohort. For a woman born in 1920, the sex ratio is calculated as number of men born in 1917–1922 divided by the number of women from the same birth cohorts. The window definition is based on the 1950 Census which shows that these are most common age differences for our birth cohorts (see Table 2.3).[15]

We compare both definitions of sex ratio measures in Figure 2.6 for the SOEP sample, finding that they are very similar except for birth cohorts around last fully drafted cohorts. If the sex ratio is defined as number of men per women by year of birth, the sex ratio

[15]We used a window of women ranging from 2 years older up to 3 years younger than men because this covers between 61% (1920) and 82% (1927) of all married couples in 1950 (see Table 2.3). For checks of the robustness to the definition of this window we also allowed for a broader window (see Figures 2.A.3.b and 2.A.4.b in the Robustness Section).

jumps back to one for the cohort of 1928. This is the first cohort of men who were not fully drafted for WWII. In contrast, the sex ratio shock is fading out more gradually when using the window definition of sex ratios.

Table 2.3: Fraction men married to women between 2 years older and up to 6 years younger, Census 1950

	100 married men belonging to birth cohorts in first column were married to a woman who was born								
	6	5	4	3	2	1	0	1	2
Birth year men	years later than husband							years earlier than husband	
1928			3.5	8.4	13.9	17.4	12.5	16.0	10.7
1927		1.4	4.3	8.3	14.6	15.7	19.1	9.4	10.8
1926	0.6	3.2	8.0	11.2	13.9	15.0	14.6	9.7	8.3
1925	1.1	3.8	6.0	11.9	12.8	15.0	15.0	11.1	7.3
1924	2.2	4.9	6.4	11.3	17.4	17.3	14.2	8.5	5.0
1923	4.3	6.3	9.6	12.2	14.3	15.6	12.1	7.3	6.7
1922	4.8	7.6	9.8	12.4	14.4	13.6	11.4	7.8	5.1
1921	4.8	8.1	10.8	13.4	14.5	14.1	13.0	5.7	3.2
1920	6.4	8.2	11.5	11.4	13.9	15.1	11.7	5.7	2.7

Source: Niehuss (2001, p. 345)

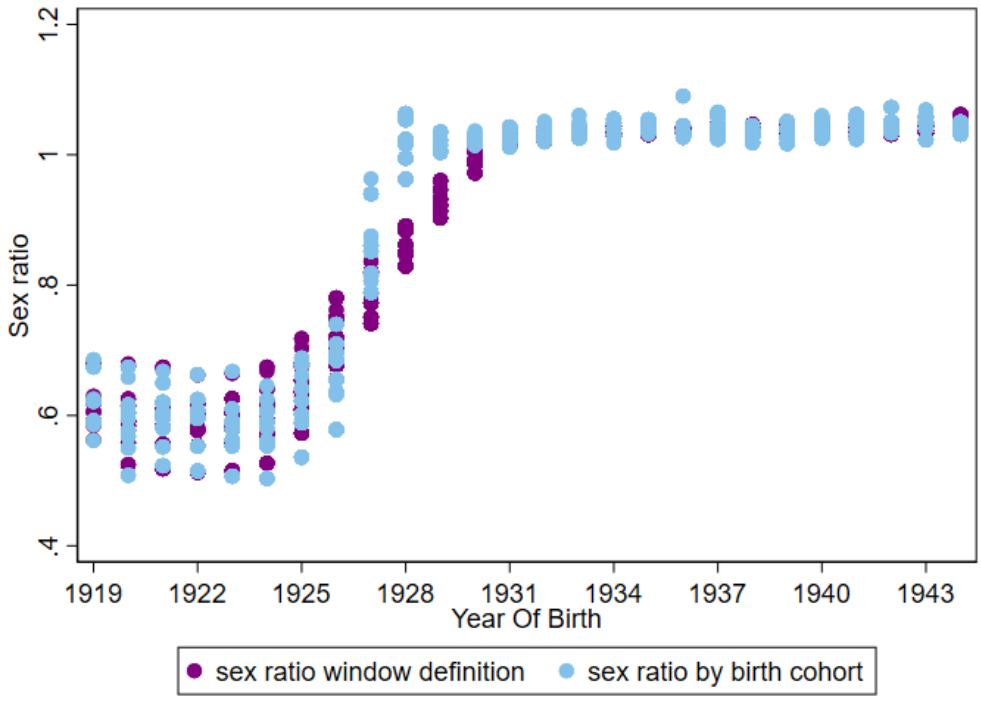
To estimate impact of unbalanced sex ratios we set up a general linear OLS model that links the outcome of interest with the sex ratio in 1946 and with a set of covariates.

$$Y_i = \alpha + \beta_1 (sexratio)_{cm} + \gamma X_i + c + m + \epsilon_i \tag{2.1}$$

In this specification, α is the constant term, X_i is a set of covariates, and ϵ_i is an error term. $(sexratio)_{cm}$ denotes the sex ratio of an individual belonging to birth cohort c in residence m. Since we have panel data, we can control for both age and cohort effects by keeping the target ages of our outcomes fixed and including 5-year cohort dummies denoted by the parameter c. The parameter m denotes a state of residence fixed effect, controlling for all time-constant differences between German regions. In all regressions, we control for individual's socioeconomic status by including a dummy that takes the value one if the father and the mother had upper secondary education, and is zero otherwise. Additionally, we control for years of schooling.[16] Standard errors are clustered on the

[16]Due to potential endogeneity issues associated with education, we ran all specifications also without

Figure 2.6: Comparison of sex ratios using +3/-2 birth cohorts to sex ratios with men and women from same birth cohorts



state of residence level (see Bertrand, Duflo, & Mullainathan, 2004).

While the sex ratio in 1946 is constant over time, its impact may change over different stages of the life cycle. To obtain life cycle profiles of fertility, we separately estimate equation (2.1) at different ages. Starting at age 20, we provide estimates for 2-3 year intervals until age 50 which we take as age of completed fertility. Estimated coefficients represent age-specific impact of a unit-change in sex ratio in 1946 on the outcome of interest. While result tables will show the original estimated coefficients, we will interpret them as the impact that a reduction in the sex ratio from 10 men in every 10 women by 1 man to 9 men in every 10 women has on outcomes of interest.

2.6 Results

The following Section provides results from estimating impacts of sex ratio in 1946 on fertility and marriage. To obtain the effects of interest at different stages over the life cycle, we run equation (2.1) for age 20, 23, 25, 28, 30, 33, 35, 38, 40, 43, 45, 48, and 50.

education. The results remain broadly unchanged and are available upon request.

Table 2.4 displays the coefficients and standard errors of the life cycle profiles for fertility and marriage outcomes.

2.6.1 Life cycle patterns of fertility and marriage

Our main interest lies in how imbalanced sex ratios influence fertility over the life cycle. Our hypothesis is that in the short run women with low sex ratios face fewer potential partners and have less children, leading to a fertility decline (Brainerd (2017) for post-WWII-Russia). We first analyze total fertility.[17] As illustrated by Figure 2.7a, we do find a significant reduction for women in their 20ies who face low sex ratios. At age 28, a decrease in sex ratio in 1946 from 10 men to 9 men for every 10 women decreases the number of children by 0.144, corresponding to an 11.5% reduction in number of children at this age-specific mean. However, the negative effects of a decrease in sex ratio on fertility at young ages turns into a positive albeit insignificant effect at age 35 and until completed fertility. This finding documents an interesting pattern: even though women exposed to a low sex ratio have fewer children when we evaluate their fertility at younger ages, they catch up with their peers who face no gender imbalance at any age later than 35 and when fertility is completed. It is important at which point in time long-term effects of unbalanced sex ratios are evaluated.

One possible pathway linking sex ratios and fertility is marriage: Most children (91.2%) are born within a marriage. The increased scarcity of men as the sex ratio goes down implies that marriage unions will be more difficult to finalize so we expect an increase in age at first marriage and a lower probability of being married at younger ages. This delay in marriage age will be exacerbated by war deployment and post-war occupation. Figure 2.7b shows impacts of unbalanced sex ratios on probability of being married. We find that reductions in sex ratios result in statistically significant decreases in the probability of being married at ages 20-38. The estimated coefficient at age 23 is 1.176, suggesting that a reduction from 10 to 9 men for every 10 women decreases the probability of being married at age 23 by about 11.8 percentage points. Given that 50% of women in our sample are married at this age, this corresponds to a reduction in marriage probability of 23.5%. The effect decreases with age, but remains statistically significant at the 5% level

[17] Number of kids woman has given birth to at given target ages includes women who are childless.

until age 38. At age 38, a reduction of 1 man for every 10 women still leads to a 3.5% lower likelihood of being married. From age 40, effects are statistically not significant anymore. This suggests that the consequences of an unbalanced sex ratio on the marriage market phase out with age and are offset in the long-run.[18] Shifts in marriage patterns from low sex ratios seem to have consequences for women's fertility pattern. As low sex ratios lead to a lower probability of early age marriage, we expect that fertility is delayed due to the scarcity of men at younger ages.

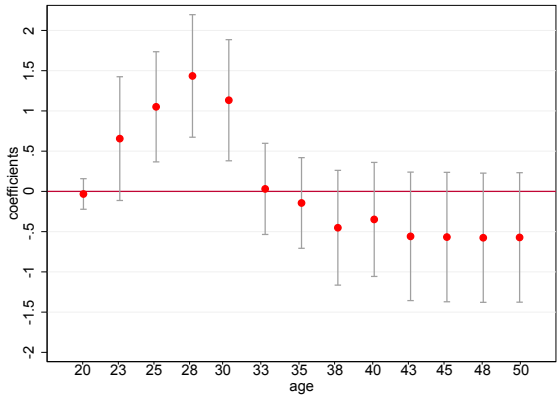
Another important driver of total fertility is the relationship between the extensive and intensive margin (Aaronson et al., 2014; Baudin et al., 2015), the share of women remaining childless and number of children among women with children. Hence, we look at the probability of remaining childless – the extensive margin of fertility – over the life cycle. Figure 2.7c shows that a decline in sex ratios significantly decreases the probability of having a child at ages 25-30. At age 25, we obtain an estimated coefficient of 0.521, suggesting that a reduction from 10 men to 9 men per 10 women reduces the probability of having a child by about 5.2 percentage points. This effect remains statistically significant at ages 28 and 30. At completed fertility, women facing a lower sex ratio still have a lower probability to have children than women facing a higher sex ratio.

What about the intensive margin? Figure 2.7d displays sex ratio effects on fertility excluding childless women. We find a similar pattern in estimated coefficients over target ages as for total fertility (Figure 2.7a). At age 28, a reduction in the sex ratio of 1 man for every 10 women significantly decreases the number of children by about 0.146, a 10.2% reduction. From age 33, women facing a low sex ratio not only catch up, but start to significantly outperform women with higher sex ratios in number of kids. This overcompensation remains statistically significant at the 5% level until age 50, implying a 5.7% higher number of children at completed fertility. Thus, for women facing low sex ratios the higher probability of being childless is overcompensated by having more children.

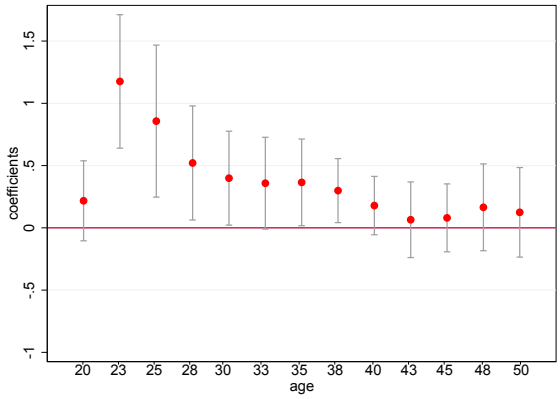
We find that women facing low sex ratios have a higher probability to be unmarried and childless in their 20s, but they catch up. This is supported by results for regressions of age at first/last birth on sex ratios in 1946 and 1950 displayed in Table 2.A.1. We find

[18] We found no significant effect of sex ratios on divorce rates for women. Results are available on request.

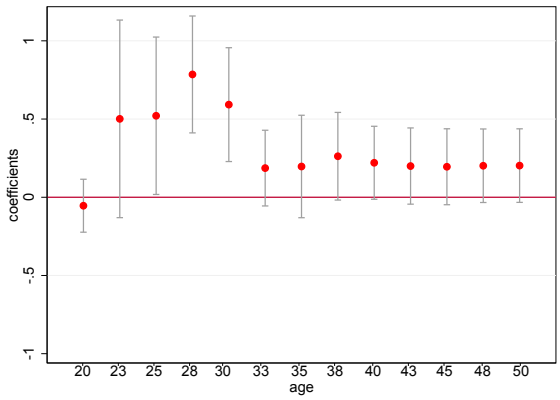
Figure 2.7: Coefficients and 95% CI for the impact of unbalanced sex ratios in 1946 on total fertility and marriage over the life cycle



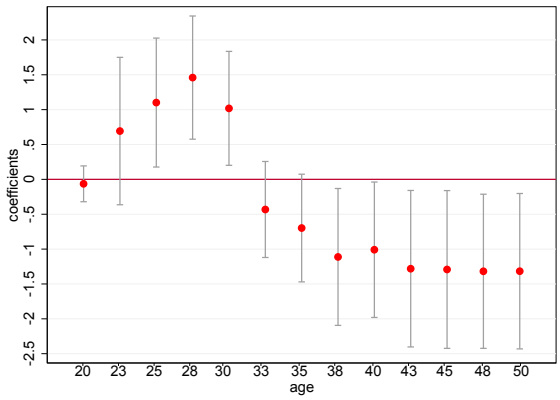
(a) total fertility



(b) probability of being married



(c) extensive margin of fertility



(d) intensive margin of fertility

Note: Estimated age-specific coefficients of sex ratio in 1946 and 95% confidence intervals are plotted against age.

that a decrease in sex ratios of 1946 by 1 man for every 10 women significantly increases mother's age at first birth by more than 6 months and age at last birth by about 8.7 months. These numbers not only suggest that women facing lower sex ratios in 1946 postpone their childbearing period. It also implies age at last birth increases even more than age at first birth, and that the childbearing period is not shorter, but longer. Column (5) in Table 2.4 confirms this finding by showing that a decrease in the sex ratio decreases the probability of having a child below age 5 at ages 23 to 28. At age 23, the coefficient is statistically significant at the 10% level. In contrast, at ages 33 and 35 a reduction in the sex ratio significantly increases the probability of having a child below age 5.

The previous analysis suggests that imbalances in sex ratios predict a decrease in extensive margin of fertility at younger ages and a catching up effect in fertility along the

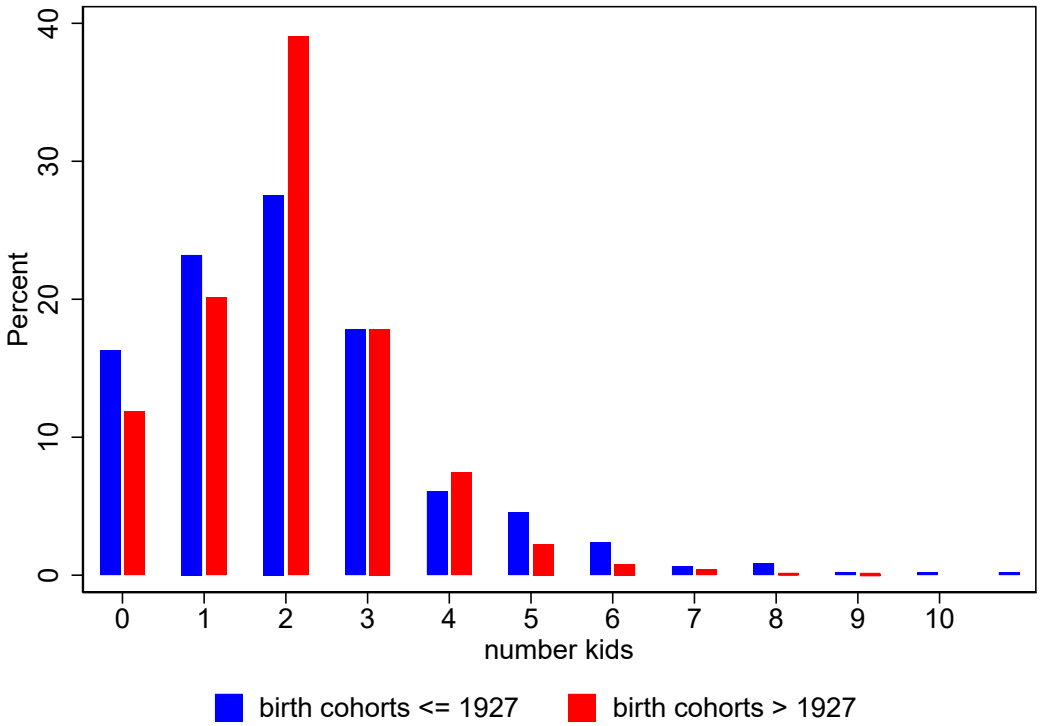
Table 2.4: Regression results of the impact of imbalances in sex ratios on fertility and marriage outcomes

	(1) Total fertility	(2) Probability first marriage	(3) Extensive margin of fertility	(4) Intensive margin of fertility	(5) Intensive margin & kid below age 5
Target age	Coef	Coef	Coef	Coef	Coef
Age 20	-0.032	0.217	-0.054	-0.064	-0.088
	(0.082)	(0.139)	(0.073)	(0.111)	(0.100)
Age 23	0.656*	1.176***	0.501	0.693	0.618
	(0.333)	(0.232)	(0.274)	(0.458)	(0.347)
Age 25	1.051***	0.857**	0.521**	1.102**	0.575*
	(0.297)	(0.265)	(0.218)	(0.401)	(0.271)
Age 28	1.435***	0.521**	0.785***	1.460***	0.493
	(0.330)	(0.199)	(0.162)	(0.383)	(0.271)
Age 30	1.133***	0.399**	0.592***	1.018**	0.416
	(0.326)	(0.164)	(0.158)	(0.354)	(0.241)
Age 33	0.031	0.358*	0.186	-0.432	-0.494**
	(0.246)	(0.160)	(0.105)	(0.299)	(0.209)
Age 35	-0.144	0.365**	0.197	-0.698*	-0.724***
	(0.244)	(0.151)	(0.142)	(0.335)	(0.216)
Age 38	-0.451	0.299**	0.262*	-1.113**	-0.542
	(0.309)	(0.111)	(0.121)	(0.426)	(0.298)
Age 40	-0.348	0.179	0.220*	-1.009**	-0.226
	(0.307)	(0.102)	(0.101)	(0.421)	(0.205)
Age 43	-0.559	0.065	0.199*	-1.281**	-0.149
	(0.346)	(0.132)	(0.106)	(0.487)	(0.138)
Age 45	-0.567	0.080	0.195	-1.292**	-0.332*
	(0.349)	(0.118)	(0.105)	(0.491)	(0.159)
Age 48	-0.576	0.165	0.201*	-1.318**	-0.046
	(0.348)	(0.151)	(0.102)	(0.479)	(0.078)
Age 50	-0.572	0.125	0.202*	-1.317**	-0.004
	(0.349)	(0.156)	(0.102)	(0.483)	(0.006)

Standard errors are clustered on the state of residence level, *** p<0.01, ** p<0.05, *p<0.1. OLS regressions at target ages. Controls: 5-year birth cohort fixed effects, state of residence fixed effects, father/mother obtained upper secondary education, years of schooling. The estimation sample in Column (3) is restricted to women's first marriage only. Number of observations: Columns (1), (3): N_{age20}-N_{age40}: 2,426; N_{age43}: 2,414; N_{age45}: 2,402; N_{age48}: 2,372; N_{age50}: 2,345; Columns (4)-(5): N_{age20}-N_{age40}: 2,118; N_{age43}: 2,106; N_{age45}: 2,094; N_{age48}: 2,069; N_{age50}: 2,043; Column varies from N_{age20}: 2,426 to N_{age50}: 2,228 by first marriage.

intensive margin for women older than 30. To illustrate the relationship between the intensive and extensive margin, we plot the distribution of total fertility by birth cohorts with balanced and low sex ratios separately (see Figure 2.8). Women with low sex ratios are overrepresented in both tails of the distribution (thus among women with very few and very many children), while women with balanced sex ratios have a more compressed fertility distribution. These distributions suggest some sort of selection mechanism. Women with strong preferences for children that were affected by the scarcity of men do get married and have many children even if they do so at later ages and, as we will see later, at the cost of obtaining lower quality matches on the marriage market. Women with weak preferences for having children tend not to marry and remain childless. This leads to a higher fraction of childless women in one tail of the distribution but also to a higher number of children in the upper tail of the distribution. Thus, in our case the intensive and extensive margin are substitutes.

Figure 2.8: Distribution of total fertility for women with low sex ratios (born before 1927) and balanced sex ratios (born after 1927)



2.6.2 Analysis of mechanisms: out-of-wedlock births and matching quality

We move to the 1% sample of the German Census in 1970 to use data not present in the SOEP (information on partners) and to re-estimate some cross-sectional results with a larger data set. While it is an advantage that 1970 is closer to WWII, so our sample is less affected by selective mortality, it is also a disadvantage, because we neither observe completed fertility nor effects at different ages. Still, we can confirm some of our results. Regression results presented in Column (1) of Table 2.5 confirm that we find that a low sex ratio significantly increases women's age at first birth. The average age at first marriage for a woman facing a completely balanced sex ratio is about 23.2. A reduction in sex ratio by 1 man to 9 men for every 10 women increases age at first marriage by about 5 months to an age of 23.7. Similarly, a highly imbalanced sex ratio delays birth of the first child (Column (2), Table 2.5). While a fully balanced sex ratio predicts an age at first birth of 24.2 years, a sex ratio of 0.9 delays this by about 4.5 months to an age of 24.7. Marriage is delayed more than childbearing so married women facing low sex ratios give birth earlier.

Table 2.5: OLS regressions of imbalances in sex ratios on age at first marriage, age at first birth and out-of-wedlock births

VARIABLES	(1) age at first marriage	(2) age first child, excl. out of wedlock	(3) Pr(out of wedlock birth)
Sex ratio (men/women), 1946	-4.439***	-3.775***	-0.043
	(0.358)	(0.368)	(0.040)
Constant	27.650***	27.952***	0.300***
	(0.721)	(0.349)	(0.029)
Observations	59,120	50,104	55,174
R-squared	0.153	0.157	0.031

Standard errors are clustered on the state of residence level, *** p<0.01, ** p<0.05, *p<0.1. OLS regressions on age at first marriage, age at giving birth for the first time (full sample and excluding out of wedlock births), probability of out of wedlock birth. Controls: 5-year birth cohort fixed effects, state of residence fixed effects, wife's and husband's education, household income.

Bethmann and Kvasnicka (2013) show that low sex ratios lead to an increase in the fraction of out-of-wedlock births in post-war Bavaria in the years right after WWII. We investigate whether out-of-wedlock births are a margin of adjustment to unbalanced sex ratios in the long-run. We regress an indicator for non-marital birth on the sex ratio and covariates. The coefficient in Column (3) of Table 2.5 is negative but insignificant. This suggests that 25 years after WWII, we do not find evidence of an increase in out-of-wedlock births in response to imbalanced sex ratios in 1946. Two reasons could explain our finding: first, while out-of-wedlock birth increased immediately after the end of WWII in 1946 to about 16.4% in West Germany, it already went down to about 11% in 1947 and further to about 8% in 1948, and thus almost reached the pre-war level of 7.8% in 1939 (see Statistisches Bundesamt, 1955, 1967). Second, while out-of-wedlock births increased by 11.6% in Bavaria in 1946, they increased by 8.7% in total West Germany (Statistisches Bundesamt, 1955).[19] Our results stand in contrast to Brainerd (2017) who documents a significant increase in out-of-wedlock births in the Soviet Union 13 years after the end of WWII. She also explains that out-of-wedlock births were explicitly encouraged by family laws in the Soviet Union to increase fertility in times of male scarcity. This was not the case in West Germany, where a conservative family structure with fathers working and mothers staying at home was encouraged (see e.g. Moeller, 1989, 1993).

Another potential mechanism is that women facing unbalanced sex ratios accept worse matches on the marriage market (Abramitzky et al., 2011). Worse matching quality in terms of husband's education may result in a higher price for child (educational) quality. Facing a trade-off between quality and quantity of children, the higher price may lead to a fertility increase along the intensive margin as women substitute out quality for quantity (Aaronson et al., 2014). At the same time the value of having children decreases, leading to higher shares of childlessness. We test this using two measures of matching quality, degree of assortative mating and age gap between spouses. Column (1) in Table 2.6 shows results from regressing years of wife's education on years of husband's education, the sex ratio in 1946, an interaction term between husband's education and the sex ratio, and on covariates. Wife's average years of education increase as the sex ratio decreases.[20]

[19] Numbers for Bavaria can be found in Bethmann and Kvasnicka (2013).

[20] For average years of husband's education, a wife facing a sex ratio of 0.7 has about 8.86 years of education while a wife with a balanced sex ratio of 1 has an average of 8.74 years of education.

Regarding matching in the education dimension, we find a positive coefficient for the interaction term. The more balanced the sex ratio, the better the educational quality of the husband or, the degree of assortative mating decreases as sex ratio imbalances increase. Consider a completely balanced sex ratio. At a sex ratio of 1, an increase in husband's education by one year is associated with an increase in wife's education by (0.179+1∗0.112) = 0.291 years. Compared to a situation with a sex ratio of 0.9, an increase in husband's education by one year is associated with an increase in wife's education by only 0.280 years. Column (2) of Table 2.6 provides results from an OLS regression of positive and negative absolute differences in the years of education by husband and wife. While not being statistically significant, both coefficients have a negative sign, thus supporting our hypothesis of a worse assortative mating for imbalances in sex ratios.

Table 2.6: Matching by education, OLS regression of wife's years of education on husband's years of education

VARIABLES	(1) wife's years education	(2) differences years education wife < husband	(3) differences years education wife > husband
Husband: years education	0.179***		
	(0.021)		
Sex ratio (men/women), 1946	-1.426***	-0.009	-0.519
	(0.263)	(0.292)	(0.410)
Husband: years educ * sex ratio, 1946	0.112***		
	(0.026)		
Constant	7.021***	2.186***	-4.484***
	(0.243)	(0.297)	(0.340)
Observations	54,977	13,336	3,762
R-squared	0.263	0.091	0.549

Standard errors are clustered on the state of residence level, *** p<0.01, ** p<0.05, *p<0.1. OLS regressions on wife's years of education, differences in education between husband and wife. Controls: 5-year birth cohort fixed effects, state of residence fixed effects, husband's year of birth fixed effects, household income. The sample consists of women in their first marriage. First marriage means that women that are separated/divorced or widowed are excluded.

Table 2.7 shows the findings from regressing spousal age gaps on sex ratios in 1946. The negative coefficient for the sex ratio in Column (1) indicates that the age gap between husbands and wives increases with a stronger imbalance of the sex ratios. Column (2) of

Table 2.7 shows how a low sex ratio influences the spousal age gap for husbands being older than their wives. If the sex ratio is equal to 1, husbands are about 6.9 years older than wives. A reduction in the sex ratio by 1 man to 9 men for every 10 women increases this age gap by 0.13 years or about 1.6 months to 7.05 years. Column (3) provides OLS results if the husband is younger than the wife. Here, the average age gap for a balanced sex ratio is 2.6 years compared to an age gap of 2.67 years if the sex ratio is 0.9. This result is broadly in line with the previous literature, for instance Brainerd (2017) who also finds adjustments to male scarcity through increased age gaps. The findings may help to explain the presumed low match quality due to the relative scarcity of men.

Table 2.7: OLS regression of spousal age gap on imbalances in sex ratios

VARIABLES	(1) spousal age gap, abs.	(2) age husband > age wife	(3) age husband < age wife
Sex ratio (men/women), 1946	-0.986***	-1.312***	-0.801***
	(0.260)	(0.330)	(0.181)
Constant	5.158***	8.227***	3.392***
	(0.177)	(0.259)	(0.171)
Observations	57,276	46,696	7,580
R-squared	0.023	0.049	0.037

Standard errors are clustered on the state of residence level, *** p<0.01, ** p<0.05, *p<0.1. OLS regressions on age differences between husband and wife; age differences if husband is older than wife; age differences if husband is younger than wife. Controls: 5-year birth cohort fixed effects, state of residence fixed effects, wife's and husband's education, household income. The sample consists of women in their first marriage. First marriage means that women that are separated/divorced or widowed are excluded.

A final potential mechanism refers to the relative bargaining power of men and women within marriage. Lower sex ratios decrease the wife's bargaining position within the household (Chiappori, Fortin, & Lacroix, 2002). Doepke and Kindermann (2017) discuss a number of studies that analyze the relationship between female bargaining power and fertility choices: While for developing countries, there is a negative relation between female bargaining power and fertility (see for instance Ashraf, Field, & Lee, 2014; Seebens, 2006; Varanasi, 2009), it seems that at least in contemporary Germany, women have higher fertility preferences than men (Hener, 2015).

2.6.3 Robustness

We next analyze the robustness of our findings on fertility. The robustness checks include several modifications of the main specification, the sample restrictions, and the way we define the sex ratio. Since we have many results, we decided to show results on total fertility and on the intensive margin of fertility. Results are available in Figures 2.A.1-2.A.4, in Appendix Section 2.A.[21] We also checked robustness with respect to the probability of being married and the extensive margin of fertility. These results are available upon request.

Our main empirical specification uses sex ratios measured in the Census one year after the end of WWII. As a first robustness check, we use the sex ratio from the 1950 Census. While the imbalance in the sex ratio was moderated slightly between 1946 and 1950 due to returning prisoners of war, sex ratios were still highly imbalanced in 1950. Almost all surviving prisoners of war had returned by the time of the 1950 Census.[22] A further alleviation of the shortage of men was hence not to be expected. As shown by Figure 2.A.1.a, the coefficients exhibit a life cycle pattern like that obtained from sex ratios in 1946. While standard errors are larger, coefficients are similar in size. At older ages, coefficients are statistically significant at the 10% level, indicating overcompensating behavior in number of children is persistent. Figure 2.A.2.a shows the same results for the intensive margin of fertility. Again, the increase in the number of children at older ages is significant at the 10% level.

We next investigate whether our findings on total fertility are driven by external shocks rather than variation in sex ratios. First, we consider the possibility that WWII had effects on women's fertility. We thus include several age-specific dummy variables in equation (2.1), taking the value one if the year when the women reaches a specific age falls within WWII. Coefficients show a very similar pattern in total fertility when controlling for these age-war dummies (Figure 2.A.1.b). Though the negative impact of a low sex ratio on total fertility is not statistically significant at younger ages, we find a catching up and significantly overcompensating impact of low sex ratios on long-term fertility. Again, Figure 2.A.2.b shows that the overcompensating effect in the intensive margin at older

[21] The estimated coefficients and standard errors for Figures 2.A.1-2.A.4 can be found in Appendix 2.B.

[22] Only a very small share was released later than 1950 (Scientific Research Service of the German Bundestag, 2011).

ages is statistically significant at the 5% level. Second, we include a measure of economic uncertainty into our regressions. Chabé-Ferret and Gobbi (2018) show that economic uncertainty (standard deviation of GDP in a certain age range) can influence fertility choices. We construct a GDP volatility measure that takes account of deviations of the GDP from its mean value up to target ages. Compared to findings on total fertility for our main specification, coefficients at young ages decrease in size while those at older ages seem to slightly increase (Figure 2.A.1.c). As shown in Figure 2.A.2.c, the coefficients for the intensive margin are positive and statistically significant at early ages and negative and statistically significant at later ages. As an alternative to controlling for specific shocks, we take account of region-specific trends by interacting calendar time and region fixed effects (Figure 2.A.1.d and 2.A.2.d). Standard errors increase, but the overall life cycle pattern of fertility remains the same, and coefficients for the intensive margin remain negative and significant at the 5% level for older ages.

In our analytic sample, the number of birth cohorts affected by unbalanced sex ratios is lower than the number of unaffected birth cohorts. To obtain more balanced treatment and control groups, we exclude the 9 youngest birth cohorts and restrict the sample to cohorts 1919-1935. As shown in Figure 2.A.1.e, despite larger confidence intervals due to lower sample size, we don't find any strong differences in size and patterns of sex ratio coefficients over the life course. Figure 2.A.2.e shows results for the intensive margin and coefficients at older ages are statistically significant on the 5% level (as in the main specification).

We further explore robustness of findings with respect to the sample definition by restricting our analysis to women who gave birth to their first child after 1945. This ensures that our results are not driven by women who already had children in 1946. An unbalanced sex ratio leads to significantly delayed fertility at earlier ages and to overcompensation in number of children at later ages (Figure 2.A.1.f and Figure 2.A.2.f). Compared to coefficients for total fertility and the intensive margin at fertility in Figure 2.7, magnitude and significance of coefficients are almost identical.

In our main specification, we used sex ratio windows with birth cohorts from up to 3 years older to 2 years younger than the female target birth cohort. In our next set of robustness checks, we investigate whether the results still hold when we change

the definition of the sex ratio window. First, we use the fraction men/women from the same birth cohort as a measure of the sex ratio. Figures 2.A.3.a and 2.A.4.a display the estimation results from the adaption how sex ratios are measured for total fertility and the intensive margin of fertility. The overall pattern is like in our main specification; the size of the coefficients goes down, but the standard errors are considerably smaller as well. In contrast to our main specification, the coefficients are statistically significant on the 5% level at later ages. Second, we change the definition of our window to include men up to two years younger and four (instead of 3) years older. As indicated by Figures 2.A.3.b and 2.A.4.b, this definition leads to more smoothing across cohorts and thus to larger standard errors, but our main results still hold. For the intensive margin of fertility, the coefficients remain negative and significant at the 10% level at older ages.

As discussed in Section 2.3.2, the sex ratios suddenly drop around the 1928 birth cohort due to the drafting pattern for men in WWII. We therefore compute a dummy variable taking the values one if a woman was born before 1928 and is zero otherwise. By using this measure in Equation (2.1) instead, Figures 2.A.3.c and 2.A.4.c indicate that, again, our main finding are robust.

We finally investigate whether age-specific patterns in our main fertility outcomes and for marriage are driven by how we define age groups, namely as single age years. Instead of estimating Equation (2.1) for single ages over the life cycle, we pool earlier ages (age 20-33) and later ages (age 34-50). Table 2.A.2 displays estimation results. Estimated coefficients for the younger age group have positive signs and are statistically significant at the 1% level. A lower 1946 sex ratio leads to a higher number of children but a lower probability of having any kids and a lower marriage probability.

Since in Germany, childbearing is associated with a high probability to leave the labor market or go from full-time to part-time employment, our results also have strong implications for the labor market. Indeed, we see the fertility patterns of women facing low sex ratios reflected in their labor market behavior. However, documenting these effects is beyond the scope of this paper and will be investigated in future work.

2.6.4 Results for East Germany

Finally, we briefly compare our findings for women from West Germany with those from East Germany. As outlined earlier, in 1949 Germany was divided into two parts: the three western occupation zones formed the Federal Republic of Germany ("West Germany") and the Soviet zone became the German Democratic Republic ("East Germany"). This separation did not only imply two distinct political systems, - a Western Democracy in the West and a socialist state in the East – but also implied two different regimes with respect to family policies.

In the East, the goal of family policy was to mobilize as many women as possible for full-time work. Similar to the Soviet Union and its satellite states, the GDR government implemented a strong pro-natalist family policy, aiming at increasing birth rates (Leitner, Ostner, & Schmitt, 2008).[23] Meanwhile, West Germany propagated a traditional family model based on marriage where men had the role of being the main breadwinner and in which women were encouraged to become full-time mothers and to not participate in the labor force (see e.g. Moeller, 1989, 1993; Pfau-Effinger, 2004).[24]

To highlight the differences in fertility patterns between West Germany and East Germany, we repeat our empirical analysis on sex ratio effects using SOEP data.[25] Figure 2.9a represents the estimated coefficients of an imbalance in sex ratio in 1946 on the extensive margin of fertility. The estimated coefficients do not follow a specific pattern as in West Germany (see Figure 2.7c). They are close to zero and insignificant in a statistical sense, implying that East German women from affected cohorts are not more likely to be childless than women born after 1928. The estimated coefficients over the life course for the intensive margin of fertility are illustrated in Figure 2.9b.

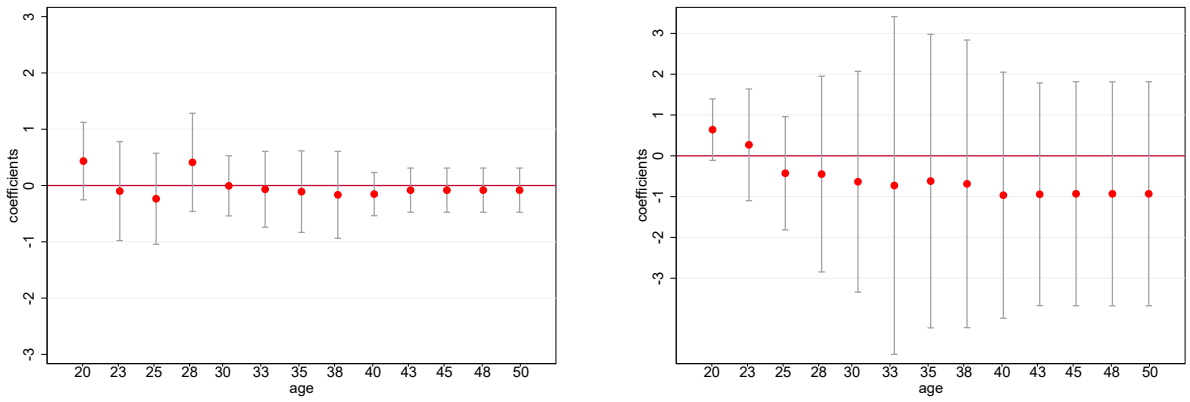
Clearly, we do not find a significant reduction for East German women in their 20ies

[23] For instance, families received loans from the government that could be amortized through having children and families with many children were privileged in the distribution of housing (Kröhnert, 2008).

[24] Through taxation and various other benefits marriage also received strong institutional support in West Germany (see e.g. Cooke, 2006; Leitner et al., 2008). In contrast, East German policies encouraged women to be economically independent from their partners. This however did not bring about economic independence, but rather resulted in relocation of dependence towards the state and the employer (Leitner et al., 2008). The East German state also established a comprehensive child care infrastructure to foster and enable female labor force participation (Zimmerman, 1993). For an comparative overviews on family policies in East and West Germany see e.g. Leitner et al. (2008) or Zimmerman (1993).

[25] In this analysis, we only present the intensive and extensive margin of fertility for East Germany. The pattern on marriage and total fertility are available upon request.

Figure 2.9: Coefficients and 95% CI for the impact of unbalanced sex ratios in 1946 on the extensive and intensive margins of fertility over the life cycle in East Germany



(a) extensive margin of fertility (b) intensive margin of fertility

Note: Estimated age-specific coefficients of sex ratio in 1946 and 95% confidence intervals are plotted against age.

who face low sex ratios regarding the number of children. Instead, this group of women seems to increase the number of children already at an early age (around age 25) and continues to have more children throughout the life cycle. However, all the estimated coefficients are statistically insignificant. Overall, the findings on fertility suggest that the strong imbalance in sex ratios had different consequences in East and West Germany. A rationale for this finding is the different institutional contexts as well as differences in the promoted role of women in the society in these two countries. While in West Germany the role of mother and housewife was promoted (Moeller, 1989, 1993), marital status was decisive (Heineman, 1999) and non-marital birth was stigmatized, single-motherhood and non-marital birth was more socially accepted in East Germany (Federal Institute for Population Research, 2018b).[26] This implies less selection into marriage and childbearing along female fertility preferences in East Germany.

[26]The share of non-marital births in East Germany was constantly higher than in West Germany from 1946 onwards (Federal Institute for Population Research, 2018b). In addition, divorce rates were higher in East Germany (Federal Institute for Population Research, 2018a). Heineman (1999, p. 13) notes that *"the Federal Republic promoted gender as a central social category; the Democratic Republic did not"* and characterizes the Democratic Republic state as a system where *"Distinct gender roles beyond those strictly tied to biological difference were understood as derivative"* (Heineman, 1999, p.14).

2.7 Conclusions

In this study, we investigate impacts of highly unbalanced sex ratios on fertility patterns over the life cycle of West German women. We argue that WWII induced an exogenous shock in the ratio of men to women in the German population which led to permanent imbalances in the sex ratios and long-term consequences on fertility behavior of women from affected birth cohorts.

To investigate our hypothesis, we combine individual level Census records and survey data. We show that women facing low sex ratios significantly delay their fertility in the short-run. In the long-run, these women exhibit a catching up and even overcompensating fertility behavior compared to women not affected by such a shock. We find that highly unbalanced sex ratios lead to a higher number of children at later ages as well as when fertility has been completed. While this finding seems to be surprising, it can be rationalized by the women's adjustments along different margins over the life cycle: They postpone first marriage and exhibit a substitutional relationship between the intensive and the extensive margin of fertility; they moreover extend their childbearing period; and accept lower quality of matches on the marriage markets.

A demographic shock that leads to unbalanced sex ratios has far-reaching consequences for both men and women. An analysis of only short-term effects on fertility would miss a significant piece of evidence. Instead, a life course perspective is required to obtain a comprehensive view on overall social and economic consequences. The revealed margins of adjustments may offer scope for policies that have attenuating and antagonizing effects.

2.A Main Appendix

Table 2.A.1: Regression results on the impact of sex ratios in 1946 and 1950 on women's age at birth of the first kid and the last kid

VARIABLES	(1) Mother's Age at Birth of		(2) Mother's Age at Birth of	
	first kid	last kid	first kid	last kid
Sex ratio (men/women), 1946	-5.305*	-7.227***		
	(2.499)	(2.068)		
Sex ratio (men/women), 1950			-6.165	6.353
			(4.322)	(3.537)
Constant	24.687***	31.900***	26.029***	32.137***
	(1.649)	(0.914)	(3.266)	(2.373)
Observations	2,037	2,038	2,037	2,038
R-squared	0.094	0.074	0.093	0.072

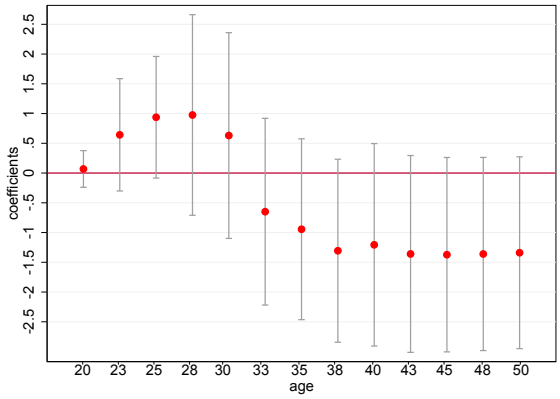
Standard errors are clustered on the state of residence level, *** p<0.01, ** p<0.05, * p<0.1. OLS regressions at target ages. Controls: 5-year birth cohort fixed effects, state of residence fixed effects, father/mother obtained upper secondary education, years of schooling.

Table 2.A.2: Regression results on the impact of sex ratios in 1946 on fertility and marriage outcomes by women's age at birth, younger versus older ages

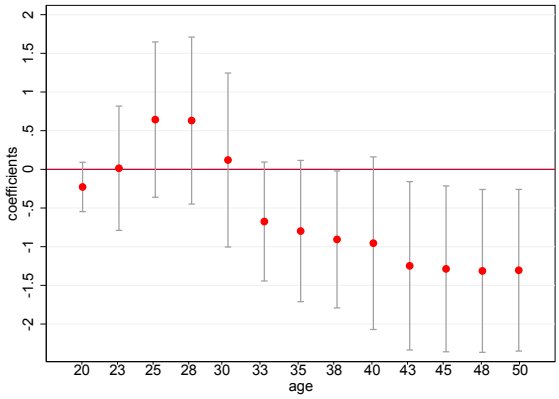
	(1) total fertility		(2) extensive margin fertility		(3) intensive margin fertility		(4) Probability marriage	
	age 20-33	age 34-50	age 20-33	age 34-50	age 20-33	age 34-50	age 20-33	age 34-50
sex ratio (men/women), 1946	0.734***	-0.439	0.438***	0.209*	0.656**	-1.115***	0.606***	0.181*
	(0.208)	(0.305)	(0.100)	(0.110)	(0.282)	(0.430)	(0.120)	(0.102)
birth cohort 1919-1924	-0.019	-0.119	0.038	0.028	-0.036	-0.251	0.062	0.007
	(0.143)	(0.139)	(0.067)	(0.055)	(0.169)	(0.211)	(0.078)	(0.062)
birth cohort 1925-1929	-0.126	-0.043	-0.050	-0.027	-0.111	-0.005	0.013	0.034
	(0.109)	(0.147)	(0.044)	(0.043)	(0.131)	(0.173)	(0.039)	(0.030)
birth cohort 1930-1934	-0.143***	0.168**	-0.076***	-0.030	-0.126**	0.275***	-0.034*	0.042*
	(0.043)	(0.065)	(0.016)	(0.023)	(0.057)	(0.071)	(0.018)	(0.024)
birth cohort 1935-1939	-0.001	0.235***	-0.031**	0.004	-0.005	0.259***	-0.032	0.028
	(0.032)	(0.048)	(0.013)	(0.006)	(0.041)	(0.066)	(0.023)	(0.020)
father has high school degree	0.088	0.059	0.027	0.032	0.076	0.004	0.003	-0.032
	(0.089)	(0.130)	(0.036)	(0.039)	(0.082)	(0.076)	(0.020)	(0.049)
mother has high school degree	-0.101	0.277	-0.061	0.006	-0.087	0.358***	-0.012	0.052
	(0.150)	(0.191)	(0.067)	(0.066)	(0.125)	(0.134)	(0.056)	(0.047)
years schooling	-0.087***	-0.083***	-0.039***	-0.023***	-0.079***	-0.045***	-0.034***	-0.017***
	(0.008)	(0.018)	(0.003)	(0.004)	(0.007)	(0.015)	(0.004)	(0.004)
Constant	1.183***	3.127***	0.615***	1.019***	1.196***	3.385***	0.452***	0.871***
	(0.206)	(0.411)	(0.104)	(0.113)	(0.309)	(0.490)	(0.144)	(0.113)
NxT	33,964	40,907	33,964	40,907	29,652	35,689	33,698	39,454
N	2,426	2,426	2,426	2,426	2,118	2,118	2,426	2,372

Standard errors are clustered on the state of residence level, *** p<0.01, ** p<0.05, * p<0.1. Random effects regression controlling for all variables presented in the Table. Reference categories: birth cohorts 1940-1944, state of residence Bremen. Models include dummy variables for the following German states: Schleswig-Holstein, Hamburg, Lower Saxony, North Rhine-Westphalia, Hesse, Rhineland-Palatinate, Baden-Wuerttemberg, and Bavaria.

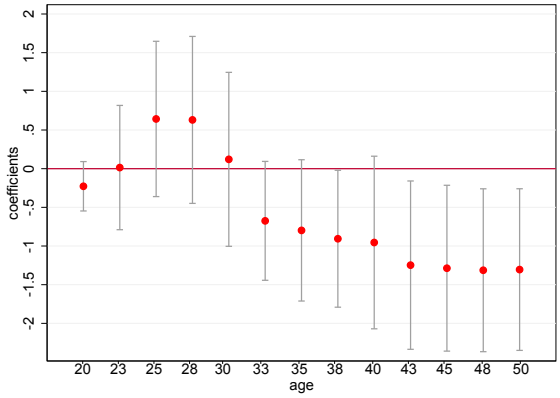
Figure 2.A.1: Regression results of the impact of imbalances in sex ratios on total fertility - Robustness checks I



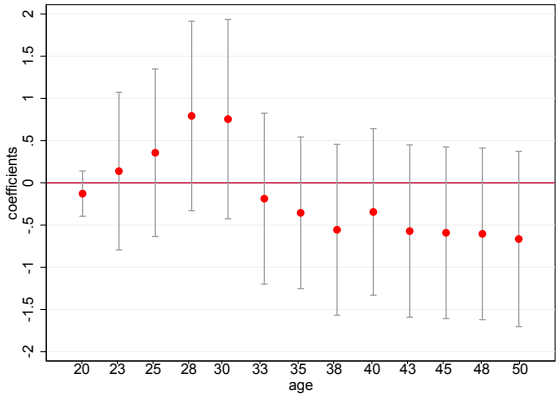
(a) sex ratio in 1950



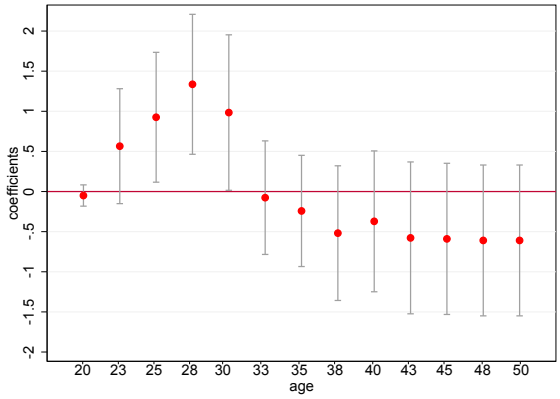
(b) control for age during WWII



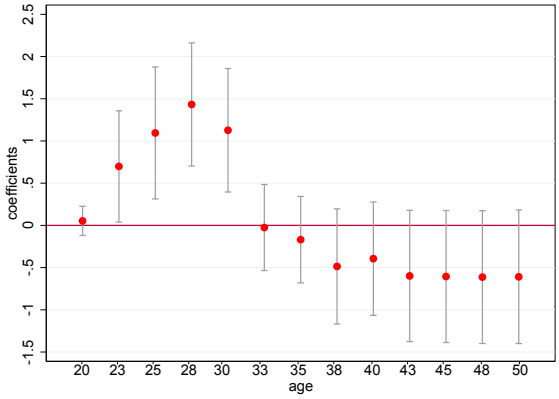
(c) control for volatility of GDP



(d) include region-specific time trends



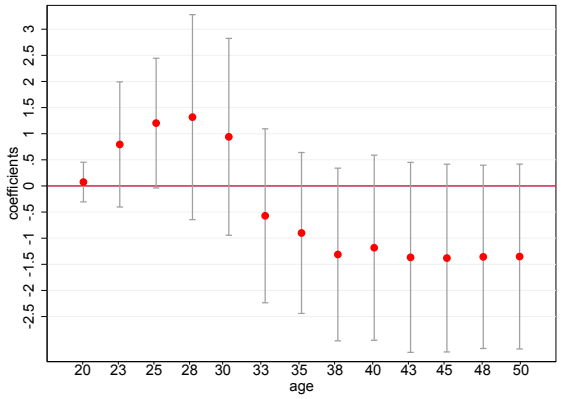
(e) only birth cohorts 1919-35



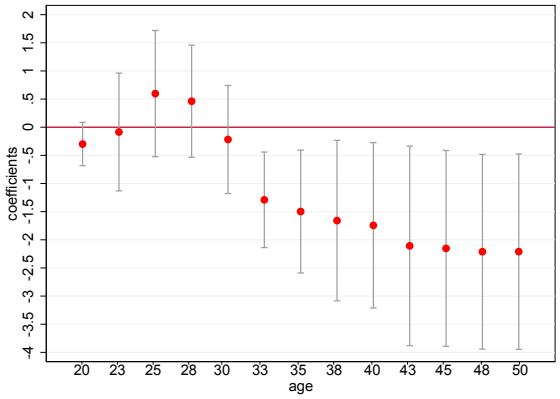
(f) no women giving birth before 1946

Note: Estimated age-specific coefficients of sex ratio in 1946 and 95% confidence intervals are plotted against age.

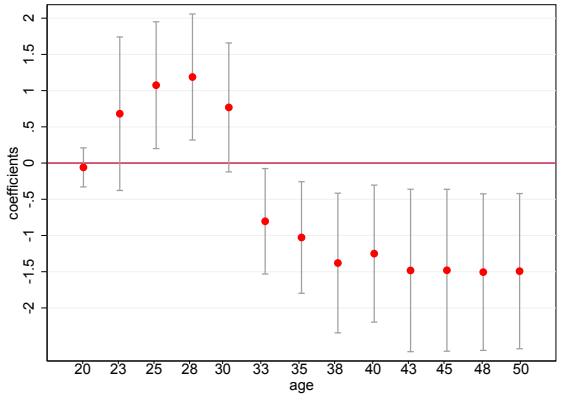
Figure 2.A.2: Regression results of the impact of imbalances in sex ratios on the intensive margin of fertility - Robustness checks I



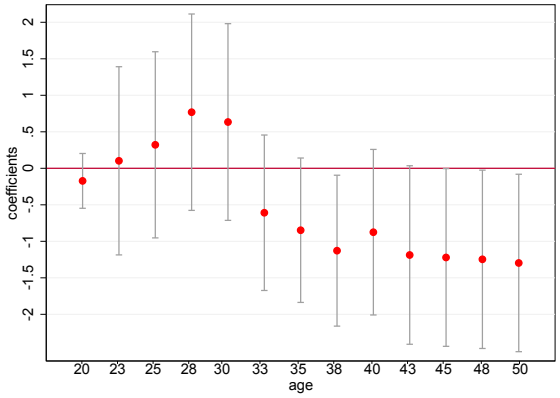
(a) sex ratio in 1950



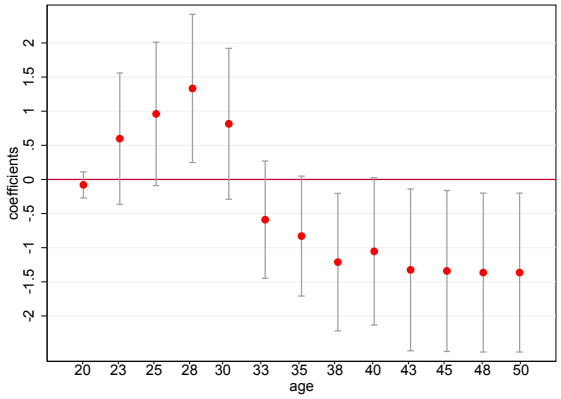
(b) control for age during WWII



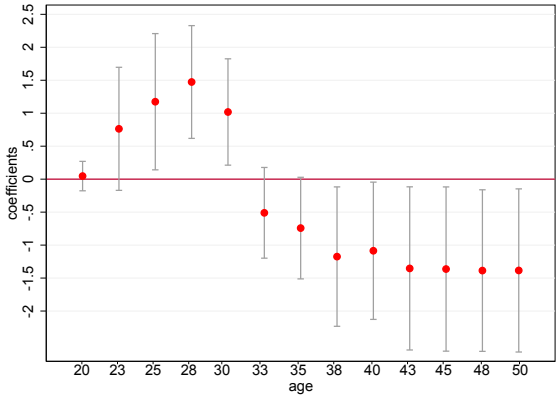
(c) control for volatility of GDP



(d) include region-specific time trends



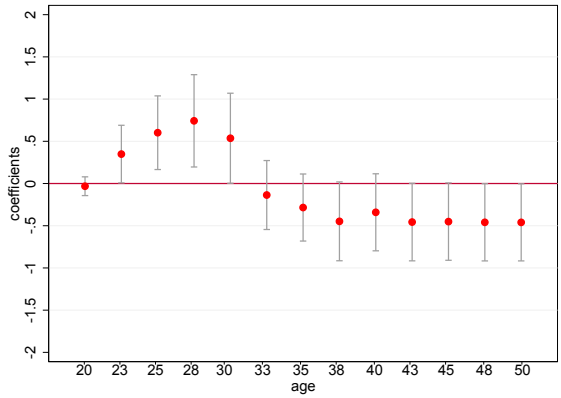
(e) only birth cohorts 1919-35



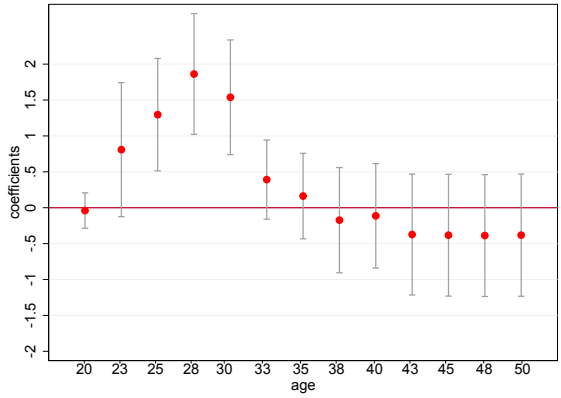
(f) no women giving birth before 1946

Note: Estimated age-specific coefficients of sex ratio in 1946 and 95% confidence intervals are plotted against age.

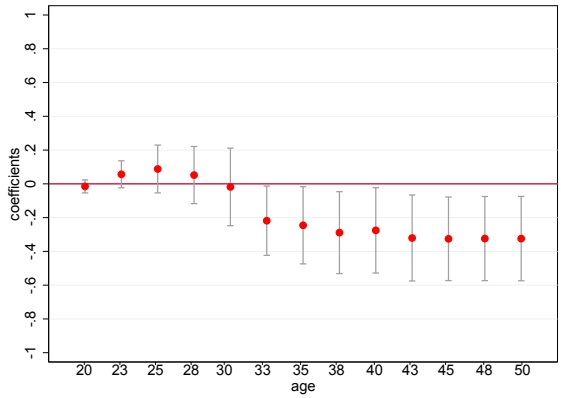
Figure 2.A.3: Regression results of the impact of imbalances in sex ratios on total fertility - Robustness checks II



(a) sex ratio window men/women same birth cohort



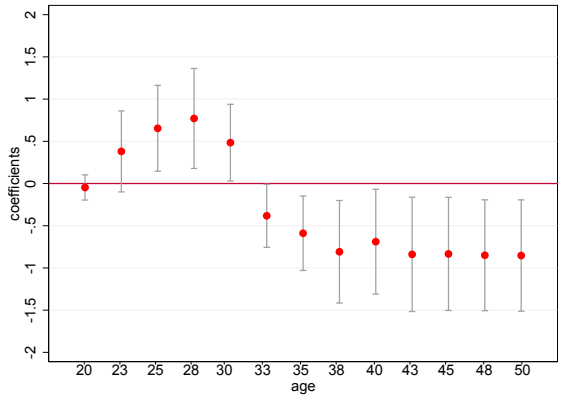
(b) sex ratio window -2/+4 years



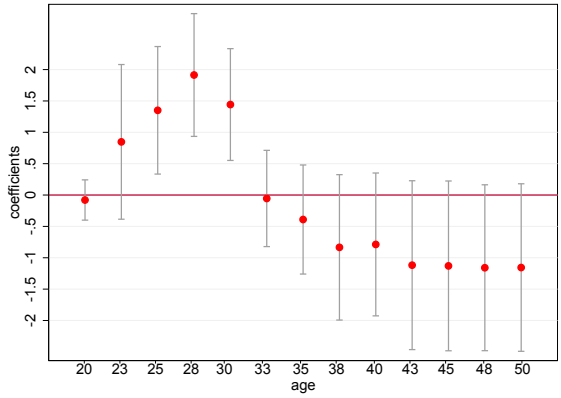
(c) dummy born before/after 1928

Note: Estimated age-specific coefficients of sex ratio in 1946 and 95% confidence intervals are plotted against age.

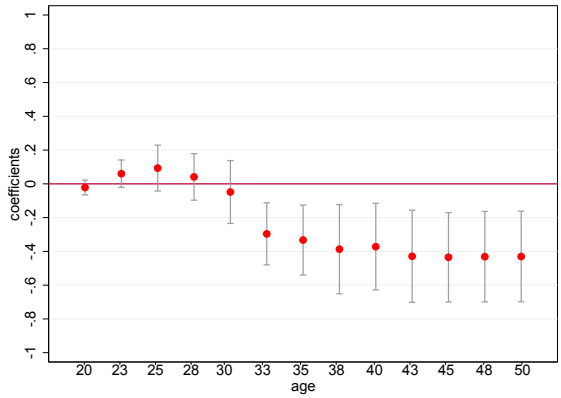
Figure 2.A.4: Regression results of the impact of imbalances in sex ratios on the intensive margin of fertility – Robustness checks II



(a) sex ratio window men/women same birth cohort



(b) sex ratio window -2/+4 years



(c) dummy born before/after 1928

Note: Estimated age-specific coefficients of sex ratio in 1946 and 95% confidence intervals are plotted against age.

2.B Online Appendix

2.B.1 Regression results for robustness checks

Table 2.B.1: Regression results of the impact of imbalances in sex ratios on total fertility – Table for Figure 2.A.1, robustness checks I

	(1) Sex ratio in 1950	(2) Control for women's age during WWII	(3) Control for volatility of GDP	(4) Region-specific time trends	(5) Only birth cohorts 1919-1935	(6) Exclude women giving birth before 1945
Target age	coef	coef	coef	coef	coef	coef
Age 20	0.068	-0.228	-0.027	-0.128	-0.050	0.053
	(0.134)	(0.138)	(0.087)	(0.116)	(0.058)	(0.075)
Age 23	0.643	0.014	0.648*	0.138	0.565	0.699**
	(0.409)	(0.348)	(0.333)	(0.405)	(0.311)	(0.285)
Age 25	0.938*	0.643	1.058***	0.357	0.926**	1.095**
	(0.443)	(0.435)	(0.287)	(0.430)	(0.351)	(0.339)
Age 28	0.976	0.631	1.167***	0.793	1.336***	1.432***
	(0.731)	(0.468)	(0.326)	(0.487)	(0.378)	(0.316)
Age 30	0.630	0.121	0.865**	0.755	0.984**	1.126***
	(0.750)	(0.488)	(0.343)	(0.512)	(0.420)	(0.317)
Age 33	-0.650	-0.675*	-0.325	-0.187	-0.076	-0.025
	(0.680)	(0.333)	(0.280)	(0.439)	(0.307)	(0.221)
Age 35	-0.944	-0.798*	-0.513*	-0.355	-0.242	-0.169
	(0.659)	(0.396)	(0.253)	(0.389)	(0.300)	(0.222)
Age 38	-1.306*	-0.906**	-0.768**	-0.556	-0.518	-0.486
	(0.667)	(0.384)	(0.332)	(0.439)	(0.364)	(0.296)
Age 40	-1.206	-0.955*	-0.625*	-0.345	-0.371	-0.394
	(0.738)	(0.484)	(0.328)	(0.428)	(0.381)	(0.291)
Age 43	-1.360*	-1.247**	-0.843*	-0.571	-0.578	-0.599
	(0.717)	(0.472)	(0.373)	(0.443)	(0.410)	(0.337)
Age 45	-1.372*	-1.287**	-0.824*	-0.591	-0.590	-0.606
	(0.708)	(0.465)	(0.369)	(0.441)	(0.408)	(0.339)
Age 48	-1.361*	-1.313**	-0.812*	-0.604	-0.609	-0.612
	(0.704)	(0.457)	(0.370)	(0.441)	(0.408)	(0.341)
Age 50	-1.339*	-1.305**	-0.820*	-0.665	-0.609	-0.609
	(0.699)	(0.453)	(0.371)	(0.450)	(0.408)	(0.343)

Standard errors are clustered on the state of residence level, *** p<0.01, ** p<0.05, *p<0.1. OLS regressions at target ages. Controls: 5-year birth cohort fixed effects, state of residence fixed effects, father/mother obtained upper secondary education, years of schooling. Number of observations: Column (1)-(4),(6): Observations vary from N=2,426 at age 20 to N=2,345 at age 50; Column (5): N=972; Column (6): Observations vary from N=2,365 at age 20 to N=2,284 at age 50.

Table 2.B.2: Regression results of the impact of imbalances in sex ratios on intensive margin of fertility – Table for Figure 2.A.2, robustness checks I

	(1) Sex ratio in 1950	(2) Control for women's age during WWII	(3) Control for volatility of GDP	(4) Region-specific time trends	(5) Only birth cohorts 1919-1935	(6) Exclude women giving birth before 1945
Target age	coef	coef	coef	coef	coef	coef
Age 20	0.074	-0.299	-0.060	-0.172	-0.079	0.047
	(0.165)	(0.167)	(0.117)	(0.163)	(0.083)	(0.097)
Age 23	0.795	-0.085	0.681	0.103	0.598	0.763*
	(0.519)	(0.453)	(0.460)	(0.559)	(0.417)	(0.405)
Age 25	1.203*	0.597	1.075**	0.322	0.961*	1.175**
	(0.538)	(0.486)	(0.380)	(0.553)	(0.455)	(0.405)
Age 28	1.317	0.462	1.188**	0.769	1.333**	1.474***
	(0.851)	(0.432)	(0.377)	(0.583)	(0.471)	(0.371)
Age 30	0.940	-0.218	0.768*	0.634	0.815	1.019**
	(0.817)	(0.416)	(0.386)	(0.584)	(0.479)	(0.350)
Age 33	-0.571	-1.290***	-0.804**	-0.609	-0.590	-0.510
	(0.722)	(0.368)	(0.316)	(0.462)	(0.373)	(0.298)
Age 35	-0.900	-1.498**	-1.028**	-0.848**	-0.830*	-0.742*
	(0.668)	(0.473)	(0.334)	(0.429)	(0.381)	(0.334)
Age 38	-1.312	-1.658**	-1.380**	-1.128**	-1.212**	-1.174**
	(0.717)	(0.618)	(0.418)	(0.448)	(0.437)	(0.458)
Age 40	-1.183	-1.743**	-1.250**	-0.875	-1.054*	-1.086**
	(0.768)	(0.637)	(0.410)	(0.492)	(0.468)	(0.451)
Age 43	-1.367	-2.108**	-1.483**	-1.187*	-1.325**	-1.353**
	(0.789)	(0.769)	(0.486)	(0.530)	(0.514)	(0.537)
Age 45	-1.381	-2.153**	-1.480**	-1.221**	-1.341**	-1.363**
	(0.779)	(0.754)	(0.485)	(0.528)	(0.511)	(0.540)
Age 48	-1.357	-2.212**	-1.506**	-1.246**	-1.365**	-1.386**
	(0.761)	(0.750)	(0.485)	(0.530)	(0.505)	(0.531)
Age 50	-1.351	-2.210**	-1.493**	-1.297**	-1.365**	-1.384**
	(0.767)	(0.753)	(0.465)	(0.527)	(0.505)	(0.536)

Standard errors are clustered on the state of residence level, *** p<0.01, ** p<0.05, *p<0.1. OLS regressions at target ages. Controls: 5-year birth cohort fixed effects, state of residence fixed effects, father/mother obtained upper secondary education, years of schooling. Number of observations: Column (1)-(4): Observations vary from N=2,118 at age 20 to N=2,043 at age 50; Column (5): N=830; Column (6): Observations vary from N=2,057 at age 20 to N=1,982 at age 50.

Table 2.B.3: Regression results of the impact of imbalances in sex ratios on total fertility – Table for Figure A.3, robustness checks II

	(1) Sex ratio as men by women from same birth cohort	(2) Sex ratios specified with a -2/+4 year window	(3) Dummy for born before/after 1928
Target age	coef	coef	coef
Age 20	-0.032	-0.040	-0.015
	(0.048)	(0.107)	(0.017)
Age 23	0.348**	0.809*	0.056
	(0.148)	(0.405)	(0.035)
Age 25	0.602**	1.296***	0.088
	(0.189)	(0.339)	(0.061)
Age 28	0.742**	1.862***	0.052
	(0.237)	(0.365)	(0.073)
Age 30	0.536**	1.538***	-0.018
	(0.231)	(0.346)	(0.100)
Age 33	-0.136	0.392	-0.218**
	(0.177)	(0.239)	(0.089)
Age 35	-0.284	0.162	-0.246**
	(0.172)	(0.258)	(0.099)
Age 38	-0.448*	-0.173	-0.289**
	(0.202)	(0.318)	(0.105)
Age 40	-0.341	-0.113	-0.275**
	(0.198)	(0.315)	(0.110)
Age 43	-0.457*	-0.373	-0.321**
	(0.199)	(0.365)	(0.110)
Age 45	-0.451*	-0.382	-0.325**
	(0.199)	(0.368)	(0.107)
Age 48	-0.459*	-0.388	-0.324**
	(0.198)	(0.368)	(0.108)
Age 50	-0.460**	-0.382	-0.324**
	(0.198)	(0.369)	(0.108)

Note: Standard errors are clustered on the state of residence level, *** p<0.01, ** p<0.05, *p<0.1. OLS regressions at target ages. Controls: 5-year birth cohort fixed effects, state of residence fixed effects, father/mother obtained upper secondary education, years of schooling. Number of observations vary from N=2,426 at age 20 to N=2,345 at age 50.

Table 2.B.4: Regression results of the impact of imbalances in sex ratios on intensive margin of fertility – Table for Figure A.4, robustness checks II

	(1) Sex ratio as men by women from same birth cohort	(2) Sex ratios specified with a -2/+4 year window	(3) Dummy for born before/after 1928
Target age	coef	coef	coef
Age 20	-0.047	-0.079	-0.022
	(0.065)	(0.139)	(0.019)
Age 23	0.380	0.848	0.060
	(0.208)	(0.535)	(0.035)
Age 25	0.654**	1.351**	0.093
	(0.221)	(0.441)	(0.059)
Age 28	0.770**	1.914***	0.041
	(0.257)	(0.424)	(0.060)
Age 30	0.484**	1.443***	-0.048
	(0.197)	(0.386)	(0.081)
Age 33	-0.382**	-0.055	-0.296***
	(0.162)	(0.333)	(0.080)
Age 35	-0.589*	-0.390	-0.333***
	(0.191)	(0.377)	(0.090)
Age 38	-0.808**	-0.834	-0.387***
	(0.263)	(0.503)	(0.115)
Age 40	-0.689**	-0.787	-0.372**
	(0.269)	(0.494)	(0.111)
Age 43	-0.839**	-1.118*	-0.429***
	(0.294)	(0.584)	(0.118)
Age 45	-0.834**	-1.130*	-0.435***
	(0.291)	(0.587)	(0.115)
Age 48	-0.850**	-1.159*	-0.431***
	(0.285)	(0.573)	(0.116)
Age 50	-0.853**	-1.156*	-0.430***
	(0.286)	(0.579)	(0.116)

Note: Standard errors are clustered on the state of residence level, *** p<0.01, ** p<0.05, *p<0.1. OLS regressions at target ages. Controls: 5-year birth cohort fixed effects, state of residence fixed effects, father/mother obtained upper secondary education, years of schooling. Number of observations vary from N=2,118 at age 20 to N=2,043 at age 50.

2.B.2 Construction of the data set on sex ratios across Census years

We collected population data by year of birth and gender on the state level for several German Censuses. However, the borders of Germany and the regional affiliation to states within Germany changed substantially over the course of the 20th century, particularly after WWII. After WWII, Germany lost all regions east of the Oder-Neisse line (Farther Pomerania, East Brandenburg, Upper Silesia, East Prussia and most of Lower Silesia). We map historic regions into todays' 16 federal states such that historic administrative regions resemble these 16 states in the best possible way. Even though our analysis focuses on West Germany, we show the mapping of pre-war states into post-unification states for all states with the exemption of Saarland.[27] Also, we discuss the variation in sex ratios for all states within the four occupational zones in Section 2.3.2 of the paper for a better illustration of the historical context.

Mapping of pre-war states/provinces

The end of WWII and the end of the Nazi Regime in Germany went along with a complete rebuilding of the political structure, which included the foundation of new federal states. While some states kept the shape and borders of pre-war states, other states were newly founded. This was mainly a consequence of the liquidation of Prussia after WWII. Some states were formed by combing two or more pre-war provinces. This is for instance the case for the state Lower Saxony as can be seen in Table 2.B.5. The state Lower Saxony consists out of the pre-war states/provinces "Provinz Hannover with Schaumburg-Lippe",[28] "Oldenburg" and "Braunschweig". However, in some cases former provinces/states were split between post-war states. In that case, an exact mapping is infeasible. This is because data on number of men and women by birth cohort is only available on the state level. Additionally, some pre-war states/provinces were split between different countries after WWII (everything East of the Oder-Neisse became a part of Poland) which also makes

[27] Saarland was not part of the four occupational zones at the time of the Census in 1946.

[28] Schaumburg-Lippe was a separate German state before WWII. Results from the census of 1939 however contain numbers for the province Hannover and Schaumburg-Lippe together.

an exact mapping infeasible.[29]

In Table 2.B.5, we describe how pre-war provinces/states relate to post-unification states. The first column contains the states, as they exist today. The second column displays how we mapped the province/states of 1939 such that correspond to the post-unification states in the best possible way. The third column describes deviations from an exact mapping or respectively, how we approximated a post-war state in case an exact mapping was not feasible.

Censuses of 1939 to 1970

Census of 1939

The Census of 1939 took place on May 17, 1939, hence 3.5 months before WWII started. Men serving their mandatory military or labor service at the time of the Census are not included in the numbers. Census results on the state level were published in several editions of the journal "Wirtschaft und Statistik" in 1940. Our procedure to map pre-war state/provinces to post-unification states is described in Section 2.B.2. All data is by year of birth.

Census of 1946

The Census was carried out within the four occupation zones and Greater Berlin in October 1946. The Census does not include the state Saarland, since it was not part of the four occupation zones when the Census took place. The state of Saarland became a part of West Germany in 1957. In the case of Baden-Württemberg (the state was founded in 1952), the numbers are based on adding up numbers for Württemberg-Baden (American Zone) + Baden (French Zone) + Württemberg-Hohenzollern (French Zone). In Figure 2.1, we give Baden-Württemberg the color of the French Zone, since two out of three former states were part of the French Zone.

For the British zone, Census data record all individuals who have attended the zone at the time of data collection. For all other occupation zones, the Census collects data on the resident population.[30] In all four occupation zones the population comprises also

[29]The website http://www.iegmaps.de/map1-0.htm provides a series of maps, which display the territorial changes of Germany across time.

[30]Members of the occupation forces are not included in the Census.

camp inmates, but excludes camp inmates of displaced person camps. While there is one comprehensive publication containing number of men and women by year of birth on the state level for that Census,[31] data on expellees by year of birth and state is not included in that publication. Data on expellees by year of birth had to be collected from publications by Statistical Offices of the states respectively the Central Statistical Office of the Soviet occupation zone. We were able to obtain these numbers for Bavaria, NRW and the states within the Soviet zone. For the other states, this information was not available.

Census of 1950 (used as robustness check in paper)

From 1950 on, East Germany and West Germany carried out separate Censuses, which also took place at different dates. The census in West Germany took place on September 13, 1950. For Hamburg and Lower Saxony data is by year of birth, for the other states data is by age cohorts. Census results were published by the Statistical Offices of the German states and hence they had to be collected from printed publications and statistical offices.

Census of 1961

The Census took place on June 6, 1961. Data is by age cohorts.

Census of 1970

The results for the Census of 1970 were kindly made available to us by the Federal Statistical Office of Germany. While the census took place on May 27, 1970, the data provided to us dates from December 31, 1970.

[31] Ausschuss der Deutschen Statistiker für die Volks- und Berufszählung 1946 (1949). Volks- und Berufszählung vom 29. Oktober 1946 in den vier Besatzungszonen und Gross-Berlin. Volkszählung: Tabellenteil. Duncker & Humblot. Berlin and Munich.

Table 2.B.5: States of Germany today and corresponding states in 1939

State today	Corresponding provinces/ states with 1939 data*	Deviations in cases where provinces/ states of 1939 cannot be mapped exactly to postunification states
Brandenburg	Provinz Mark Brandenburg	The province "Mark Brandenburg" included Ostbrandenburg, which became a part of Poland after WWII.
Mecklenburg-Vorpommern	Mecklenburg + Provinz Pommern	The province "Pommern" included Hinterpommern, which became a part of Poland after WWII.
Lower Saxony	Provinz Hannover with Schaumburg-Lippe + Oldenburg + Braunschweig	
Hesse	Provinz Hessen-Nassau + Hessen	
Saxony-Anhalt	Provinz Sachsen + Land Anhalt	
Baden-Württemberg	Baden + Württemberg with Hohenzollerische Lande	
North Rhine-Westphalia (NRW)	Provinz Westfalen with Land Lippe	The northern part of the Prussian province "Rheinprovinz" (administrative districts Aachen, Düsseldorf and Köln) also became a part of NRW after WWII. Since there is no data available for these smaller regions we don't include "Rheinprovinz" for the construction of the sex ratios for NRW.

State today	Corresponding provinces/ states with 1939 data*	Deviations in cases where provinces/ states of 1939 cannot be mapped exactly to postunification states
Rheinland-Pfalz (RLP)	Regierungsbezirk Pfalz	The state Rheinland-Pfalz in its todays borders consists of the following former regions: The southern part of the Prussian "Rheinprovinz" (administrative districts Trier and Koblenz), the western part of the Prussian province Hessen-Nassau and the Southeastern part of Hessen as well as the Regierungsbezirk Pfalz. Since there is no population data by year of birth on that regional level, we use sex ratios of the Regierungsbezirk Pfalz as a proxy for sex ratios of RLP since this region is the only one fully contained in the postwar state RLP.
Thuringia	Provinz Thüringen	
Saxony	Sachsen	
Bremen	Bremen	
Berlin	Berlin	
Hamburg	Hamburg	
Bavaria	Bayern rechts des Rheins	
Schleswig-Holstein	Provinz Schleswig-Holstein	

*Note: Names of the administrative regions /provinces in this column are the German names of the states/provinces as they are entailed in the publications containing the final census results.

2.B.3 References Census data

a) Census 1939

Statistisches Reichsamt (1940a). Weitere Teilergebnisse der Volks- und Berufszählung

vom 17.Mai 1939. Die ständige Bevölkerung nach Geschlecht, Alter und Familienstand. *Wirtschaft und Statistik*, 20(12).

Statistisches Reichsamt (1940b). Weitere Teilergebnisse der Volks- und Berufszählung vom 17.Mai 1939. Die ständige Bevölkerung nach Geschlecht, Alter und Familienstand. *Wirtschaft und Statistik*, 20(15).

Statistisches Reichsamt (1940c). Weitere Teilergebnisse der Volks- und Berufszählung vom 17.Mai 1939. Die ständige Bevölkerung nach Geschlecht, Alter und Familienstand. *Wirtschaft und Statistik*, 20(16).

Statistisches Reichsamt (1940d). Weitere Teilergebnisse der Volks- und Berufszählung vom 17.Mai 1939. Die ständige Bevölkerung nach Geschlecht, Alter und Familienstand. *Wirtschaft und Statistik*, 20(17).

Statistisches Reichsamt (1940e). Weitere Teilergebnisse der Volks- und Berufszählung vom 17.Mai 1939. Die ständige Bevölkerung nach Geschlecht, Alter und Familienstand. *Wirtschaft und Statistik*, 20(19).

b) Census 1946

Ausschuss der Deutschen Statistiker für die Volks- und Berufszählung 1946 (1949). *Volks- und Berufszählung vom 29. Oktober 1946 in den vier Besatzungszonen und Gross-Berlin. Volkszählung: Tabellenteil.* Duncker & Humblot. Berlin and Munich.

b).1 References for Figure 2.3

Bayerisches Statistisches Landesamt (1948). *Statistisches Jahrbuch für Bayern 1947*, vol. 23.

Statistisches Landesamt Nordrhein-Westfalen (1950). Landes- und Kreisergebnisse aus der Volks- und Berufszählung 1946, *Beiträge zur Statistik des Landes Nordrhein-Westfalen*, Issue 3.

Statistisches Zentralamt für die sowjetische Besatzungszone (1949). *Volks- und Berufszählung vom 29. Oktober 1946 in der sowjetischen Besatzungszone.* Band III: Landes- und Kreisstatistik, issues 1-5.

c) Census 1950 in West Germany

Bayerisches Statistisches Landesamt (1952). Volks- und Berufszählung am 13. September 1950 in Bayern. Volkszählung. I. Teil: Gliederung der Wohnbevölkerung. *Beiträge zur Statistik Bayerns*, vol. 171, 26-29.

Hessisches Statistisches Landesamt (1952). Die Struktur der Bevölkerung in Hessen. Endgültige Ergebnisse der Volkszählung vom 13. September 1950. *Beiträge zur Statistik Hessen.* Sonderreihe Volkszählung 1950, issue 3, 12-15.

Niedersächsisches Amt für Landesplanung und Statistik (1953). Die Bevölkerung am 13. September 1950 nach Altersjahren und Geschlecht. *Statistisches Jahrbuch für Niedersachsen*, 1952.

Statistisches Landesamt Baden-Württemberg (1959). Ergebnisse der Volks- und Berufszählung vom 13. September 1950. *Statistik von Baden-Württemberg*, vol. 4.1, 46-51.

Statistisches Landesamt der Freien und Hansestadt Hamburg (1953). Die Volkszählung in Hamburg am 13. September 1950. *Statistik des Hamburgischen Staates*, issue 36, 13-14.

Statistisches Landesamt Nordrhein-Westfalen Düsseldorf (1952). Die Wohnbevölkerung in Nordrhein-Westfalen nach Alter, Familienstand und Wohnort am 1. September 1939. Ergebnisse der Volkszählung vom 13. September 1950. *Beiträge zur Statistik des Landes Nordrhein-Westfalen*, Sonderreihe Volkszählung 1950, issue 3, 16-19.

Statistisches Landesamt Rheinland-Pfalz (1953). Volkszählung am 13. September 1950. Die Wohnbevölkerung in Rheinland-Pfalz, Landesergebnisse. *Statistik von Rheinland-Pfalz*, vol. 17, Bad Ems.

Statistisches Landesamt Bremen (1953). Die Volks- und Berufszählung am 13.9.1950 im Lande Bremen. *Statistische Mitteilungen aus Bremen.* Sonderheft 4.

Statistisches Landesamt Schleswig-Holstein (1953). Die Bevölkerungsstruktur in Schleswig-Holstein. Ergebnisse der Volkszählung vom 13. September 1950. *Statistik von Schleswig-Holstein*, issue 8, 4-5.

d) Census 1961 in West Germany

Statistisches Bundesamt Wiesbaden (1966). Volks- und Berufszählung vom 6. Juni 1961. Heft 4. Bevölkerung nach Alter und Familienstand, Fachserie A, Bevölkerung und Kultur. W. Kohlhammer GmbH, Stuttgart und Mainz, p.70-117.

e) Census 1970 in West Germany

Statistisches Bundesamt (2017). Bevölkerung nach Bundesländern 1970, Wiesbaden.

3 Historic sex-ratio imbalances predict female participation in the market for politicians

with Iris Grant, Iris Kesternich, and Joachim Winter[1]

This paper is published as "Historic sex-ratio imbalances predict female participation in the market for politicians" in the Journal of Economic Behavior & Organization (2018), Volume 156, p.144-165.

Abstract

We analyze the long-term effects of gender imbalances on female labor force participation, in particular in the market for politicians. We exploit variation in sex ratios – the number of men divided by the number of women in a region – across Germany induced by WWII. In the 1990 elections, women were more likely to run for office in constituencies that had relatively fewer men in 1946. We do not find a significant effect of the sex ratio on the likelihood of a woman winning the election. These results suggest that while women were more likely to run for a seat in parliament in constituencies with lower historical sex ratios, voters were not more inclined to vote for them. Voter demand effects thus do not appear to be as strong as candidate supply effects.

Keywords: Female politicians; Gender stereotypes; Occupational choice; Sex imbalance
JEL codes: J16; J24; N44

[1]The authors gratefully acknowledge financial support by KU Leuven research funds (Grant and Kesternich), the Swiss National Fund for Scientific Research through grant SNF 100018E-161598 (Steckenleiter), and Deutsche Forschungsgemeinschaft through CRC TRR 190 (Winter). We would like to thank the Federal Returning Officer's office for providing data on German elections; the staff of various statistical offices for their help in collecting the census data; Ferdinand Kramer for providing expertise on the historical background; two referees for the helpful comments and suggestions; and Kevin Kloiber and Valentin Kecht for excellent research assistance.

3.1 Introduction

Worldwide, only 22 percent of parliamentarians and 17 percent of government ministers are women.[2] In West Germany, the share of female members of parliament at the federal level has been well below 20 percent in the post-war decades. In the first federal election after the German Unification in 1990 it reached just 20.5 percent, and it is still below 37 percent (Bieber, 2013). Politicians' gender can have important consequences for policies (Beaman, Chattopadhyay, Duflo, Pande, & Topalova, 2009; Chattopadhyay & Duflo, 2004; Powley, 2006). If women are held back because of agency problems between party leaders and voters, welfare is decreased (Casas-Arce & Saiz, 2015). Further, exposure to female leaders can change perceptions about women's (lower) ability as politicians (Beaman et al., 2009). Because female politicians matter, more than 100 countries have introduced affirmative action programs for women in political office (Beaman et al., 2009).

The reasons why women are less likely to be active as members of parliament, and poorly represented in leadership positions more generally, are not well understood. A recent literature, reviewed below, stresses the role of past experiences and cultural norms as potential determinants of female labor force participation, especially in high-profile professions. Consistent with this notion, this paper highlights the fact that historical shocks to the gender balance still affect female participation in leadership professions after many decades. Our analysis shows that the drastic imbalances in the sex-ratio – defined as the number of men divided by the number of women in a region – induced in Germany by the Second World War (WWII) are strong predictors of the participation of female candidates in the first federal election after unification, more than four decades after the war ended. We argue that these post-war sex ratios exhibit exogenous variation across voting districts. Among the most affected cohorts (those who were drafted into the German army), up to 38 percent of men died but there was random variation across soldiers' home regions. The location of front lines and the allocation of prisoners of war when combat ceased also contributed to the regional variation in post-war sex ratios.

We collected data on sex ratios at the county level for the period immediately after the end of WWII and combined them with election data for the German federal parliament

[2]Numbers as of August 2016, see United Nations Women (n.d.).

(*Bundestag*). The analysis focuses on the first election after re-unification in 1990 so that all current German states can be included. We first estimate a regression of the share of female candidates from all parties in the constituencies on the 1946 sex ratio, using several specifications. Our main result is that a one standard deviation decrease in the 1946 sex ratio was associated with an increase in the share of female candidates in the 1990 federal election of 3.9 percentage points. Probit regressions of gender on the sex ratio for nominated and elected candidates indicate that these results are mainly driven by candidate supply effects since only the effect on the number of candidates, not the effect on the number of elected women, is significant. If party leaders were holding back women at the expense of votes, then we would expect more women to be elected as soon as more are nominated. Since this is not the case, we argue that party leaders' preferences are unlikely to be main explanation for the effects we find.

While long-run labor market effects of sex-ratio imbalances have been documented before, we study a shock that is relatively localized. In particular, previous papers have suggested causal channels related to culture and institutions which in Germany did not change in the post-war period at the level of the identifying variation we use, that is, voting districts.[3]

The premise of our analysis is that constituencies which experienced a more severe shortage of men relative to women after the war – in short, which had lower sex ratios – have a larger share of female nominees in the 1990 federal elections. The argument is as follows. The shortage of men right after the war made it necessary for women to step in.[4] Hence, people were exposed to women in occupations and positions that were previously dominated by men. The salience of female "role models" induced adolescent girls to raise their aspirations and aim for better jobs. Furthermore, the presence of women in high-level jobs may affect beliefs about their ability to perform well in these jobs. We present some evidence in support of this channel by showing that around the time of our elections, female labor market participation is also higher in regions with lower sex ratios. We find that the effect of unbalanced sex ratios on female nominees is larger than the effect on labor force participation (in absolute terms). Another potential channel, which we also

[3]There were of course major differences in the development of culture and institutions between East and West Germany but as we show below, this does not explain our findings.

[4]A comprehensive overview of the immediate post-war period in Germany can be found in Niehuss (2001).

discuss below, is related to the larger share of female voters, which in itself might have made parties more open towards female-oriented policies and female candidates.

The remainder of this paper is structured as follows. We review the related literature in Section 3.2. In Section 3.3, we describe the data. Section 3.4 provides a detailed discussion of the sources of variation in post-war sex ratios and of potential channels for long-lasting sex-ratio effects. Our main results are presented in Section 3.5. We discuss two potential channels through which lower sex ratios led to an increase in the share of female politicians in Section 3.6. Robustness checks are presented in Section 3.7. Section 3.8 concludes.

3.2 Related literature

There are several reasons why an under-representation of women in politics is considered to be undesirable (see, e.g., Casas-Arce and Saiz (2015)). First, women make up half of the (voting) population. A more equal representation of women at the highest level can improve the protection of their rights and interests. Second, public goods provision may be different between male and female politicians. When women are elected, they often concentrate on social policies such as education, health and childcare. These policies tend to be more beneficial towards their own sex (Bratton & Ray, 2002; Chattopadhyay & Duflo, 2004). Other positive effects of a higher share of female politicians in parliament include lower levels of corruption (Dollar, Fisman, & Gatti, 2001).

The literature identifies various reasons why fewer women occupy elected political offices. Baltrunaite, Casarico, Profeta, and Savio (2016), for instance, distinguish two channels that impact female representation: voter preferences and the parties' selection of candidates. Casas-Arce and Saiz (2015) find evidence that party leaders are less inclined to nominate female candidates even at the expense of receiving fewer votes. Additionally, they point out that women themselves might be more reluctant to enter the market for politicians. This channel is related to an emerging literature on women's reluctance to enter competitive environments, beginning with Niederle and Vesterlund (2007). Relatedly, established social norms typically attribute male characteristics to certain jobs or leadership positions (Akerlof & Kranton, 2000; Eagly & Karau, 2002). Such norms

might discourage adolescent girls from aspiring for high-level (political) positions and from making the corresponding educational choices. This premise is supported, *inter alia*, by a randomized control trial that shows that prior exposure to female leaders leads to electoral gains for women, see Beaman, Duflo, Pande, and Topalova (2012).

Besides these demand and supply factors, historically grown institutional settings can help or hinder the success of women. Matland and Studlar (1996) stress the importance of how the electoral system is organized, i.e., proportional representation versus winner-takes-all. Also, cultural and psychological aspects such as gender roles can influence female participation in the market for politicians.

Sex ratios have been widely used to study the consequence of imbalances in labor and marriage markets (Guttentag & Secord, 1983). Imbalances in sex ratios are also of interest *per se*: As the technology to control sex ratios at birth has become widely available, sex ratios are increasingly imbalanced in several countries. Such countries include India and China whose population makes up a large share of the world's population (Attané, 2006).

Prior literature has established that lower sex ratios, i.e., fewer men relative to women, imply higher female labor market participation rates. Angrist (2002) establishes this fact for second generation immigrant women in the early 20th century by using variation in immigrant sex ratios in the U.S. Acemoglu et al. (2004) use regional variation in the U.S. military mobilization for WWII to show that in states with higher mobilization rates more women had joined the workforce after the war. Even though in the U.S., unlike in Germany, the drop in the sex ratios caused by WWII was temporary rather than permanent, it has been shown to have a long-lasting impact on female labor force participation: Fernández et al. (2004) document that sons of women who entered the labor market to replace absent men were more likely to be married to working women later on. Thus, being exposed to a working mother seems to change attitudes towards working wives. Two recent papers imply that unbiased sex ratios can have long-term impacts on gender roles in the labor market. Teso (2019) documents that female-biased sex ratios caused by the slave trade have permanently altered gender roles in Africa. Grosjean and Khattar (2019) document that in areas in Australia that were more male-biased historically, people are more conservative with respect to female labor force participation, and that in these areas women work less often in high-ranking occupations.

In the analysis of sex ratios, it is difficult to establish causal effects as the presence of omitted variables or reverse causality usually cannot be ruled out. Men or women may move to regions of a country for reasons related to local marriage or labor markets that are correlated with outcomes of interest. To address this problem, variation in sex ratios outside of individual control can be exploited. As mentioned above, Angrist (2002) exploits changes in immigrant sex ratios in the US to study labor market outcomes. Abramitzky et al. (2011) use changes in sex ratios in France caused by World War I to investigate changes in mating patterns. We use a similar source of variation as considered by Acemoglu et al. (2004) and Brainerd (2017). Brainerd (2017) uses both sex ratios and outcomes measured in the first Russian postwar census of 1959 to document that the dramatic drops in sex ratios for several cohorts and regions in Russia led to lower rates of marriage and higher rates of out-of-wedlock births and divorces for women. Finally, in recent related work Kesternich, Siflinger, Smith, and Steckenleiter (2018) exploit variations in post-war sex ratios in Germany to document long-lasting effects of sex ratios on marriage and fertility patterns using individual-level data from the German Socio-Economic Panel.

3.3 Data

3.3.1 Outcome variables

In our subsequent analysis, the main outcome variable is the share of female candidates in the 1990 *Bundestag* election. As we explain below, each party selects one candidate per district in a winner-takes-all system. We then analyze whether a female candidate was elected. The unit of observation is the voting district. Voting districts in Germany are designed to roughly represent the same number of voters, and often correspond to political counties. In the 1990 elections, there were 328 districts.

Germany has a mixed election system. Voters cast two votes, one for a district candidate (this is often referred to as the "first vote" or the "direct mandate") and another for a political party (the "second vote"). The second vote determines the overall share of the parties in the parliament, which is however not relevant to the present analysis. Each party can nominate one candidate for the first vote per voting district. Additionally, independent candidates are allowed, but they typically command only very small shares

of the vote.

Our dataset, obtained from the Federal Returning Officer, covers all 2,338 candidates nominated for the first vote in the 1990 elections (Der Bundeswahlleiter, 2016b). It includes the sex and party affiliation of all candidates as well as the vote shares of all parties by voting district. We also collected data on the 1987 West-German election and on the composition of the East-German *Volkskammer* before unification.[5] This allows us to determine whether the last election's incumbent runs for office again. Finally, we analyze data from the 1994 election as well as the last pre-unification election in West Germany in 1987 as robustness checks. To obtain incumbent information for the 1987 election, we additionally collected data on the West German election of 1983. Data on these three elections was also provided by the Federal Returning Officer (Der Bundeswahlleiter, 2016a, 2017a, 2018).

Information on female labor force participation in 1997 is obtained from the INKAR database.[6] 1997 is the year closest to our election data for which we could obtain (female) employment indicators on the county level.

3.3.2 Sex ratios

With these data, we merged data on the sex ratios immediately after WWII, collected from the 1946 census which was conducted by the four occupation forces (U.S., U.K., France and the Soviet Union).[7] These data were digitalized from printed publications containing the final results of the census (Ausschuss der deutschen Statistiker für die Volks– und Berufszählung 1946, 1949a). We additionally collected data on pre-war sex ratios on the county level in order to compare pre-war and post-war sex ratios. Pre-war sex ratios stem from published results of the German census in May 1939 (Statistisches Reichsamt, 1941).

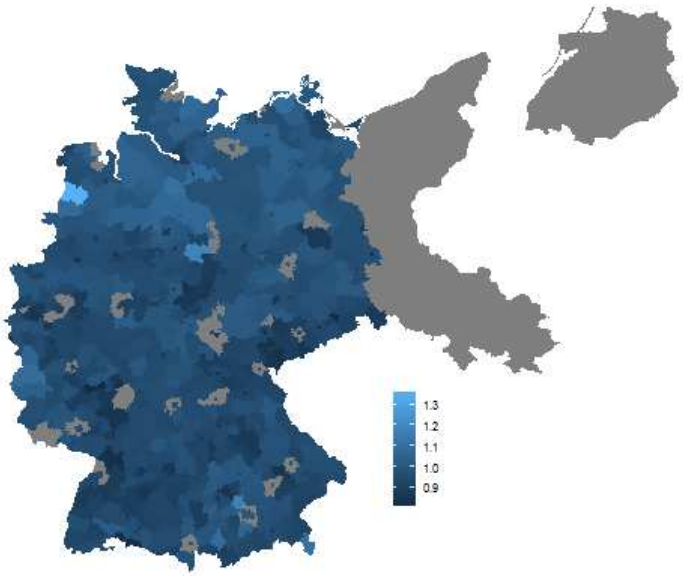
Figures 3.1a and 3.1b display the sex ratios of 1939 and 1946, respectively, for the counties situated within the present borders of Germany. The underlying map is based

[5]For details on the data sources, see Data Appendix 3.A.2.

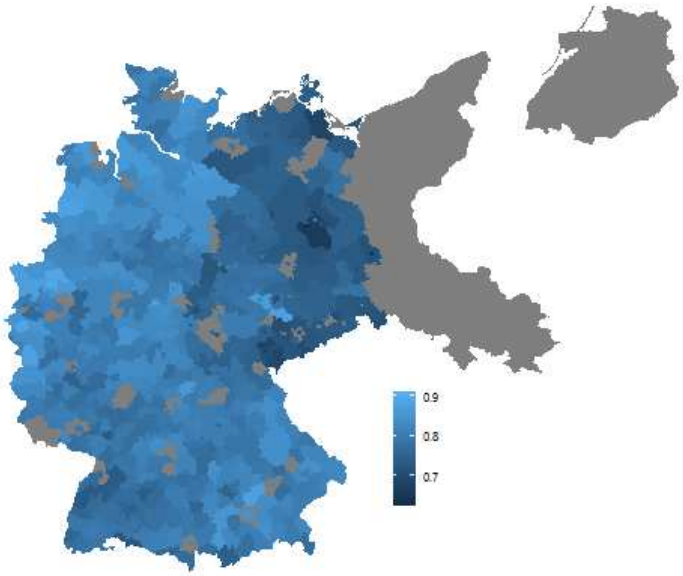
[6]Indicators and maps on spatial and urban development in Germany (INKAR), 2017 edition, are provided by the Federal Institute for Research on Building, Urban Affairs and Spatial Development (BBSR) of the Federal Office for Building and Regional Planning (BBR). The database can be accessed on http://www.inkar.de.

[7]The state of Saarland (with five voting districts in 1990) is not included in our dataset since it was not part of the four occupation zones in 1946 and no comparable census data are available.

Figure 3.1: Pre- and post-war county level sex ratio variation in Germany



(a) 1939



(b) 1946

on the county borders of the German Reich before WWII. Hence, due to administrative and territorial changes both in the 1930s and when compared to 1946, not all counties could be matched to the map. The non-matched counties are colored in grey.

Information on the numbers of men and women by year of birth is not available at the county level. However, to illustrate that the drafting pattern of the *Wehrmacht* had

a strong impact on post-war sex ratios, we additionally collected data on sex ratios in 1939 and 1946 by birth cohort at state level from various printed publications. To assess the influence of expellees on the sex ratios, we collected data on expellees at the county as well as state level. At the state level (by year of birth), we were able to obtain this information for the states in the Soviet zone, North Rhine-Westphalia and Bavaria. Data on the number of expellees in a county are available for counties in the British, American and Soviet Zone, while they are not available for counties in the French zone.[8]

We end by illustrating the fact that in Germany, the shock to sex ratios brought about by WWII was much more long-lasting than in the U.S. where returning soldiers brought sex ratios back to normal values within a few years (see Acemoglu et al. (2004)). Figure 3.2 and Figure 3.3 show sex ratios by year of birth, state, and different points in time given by census years (1946, 1950 and 1964 for East Germany, and 1946, 1950, 1961 and 1970 for West Germany). The striking drop in sex ratios is evident even in the censuses of the 1960s and 1970s. Even though these figures display considerable variation across states, we should stress that we exploit more granular variation at the level of voting districts, controlling for state fixed effects and, as a robustness check, also stratifying by East and West Germany.

[8]Detailed information on all historical data used in the analysis are given in the Data Appendix.

Figure 3.2: Sex ratios in 1946, 1950 and 1964 by state

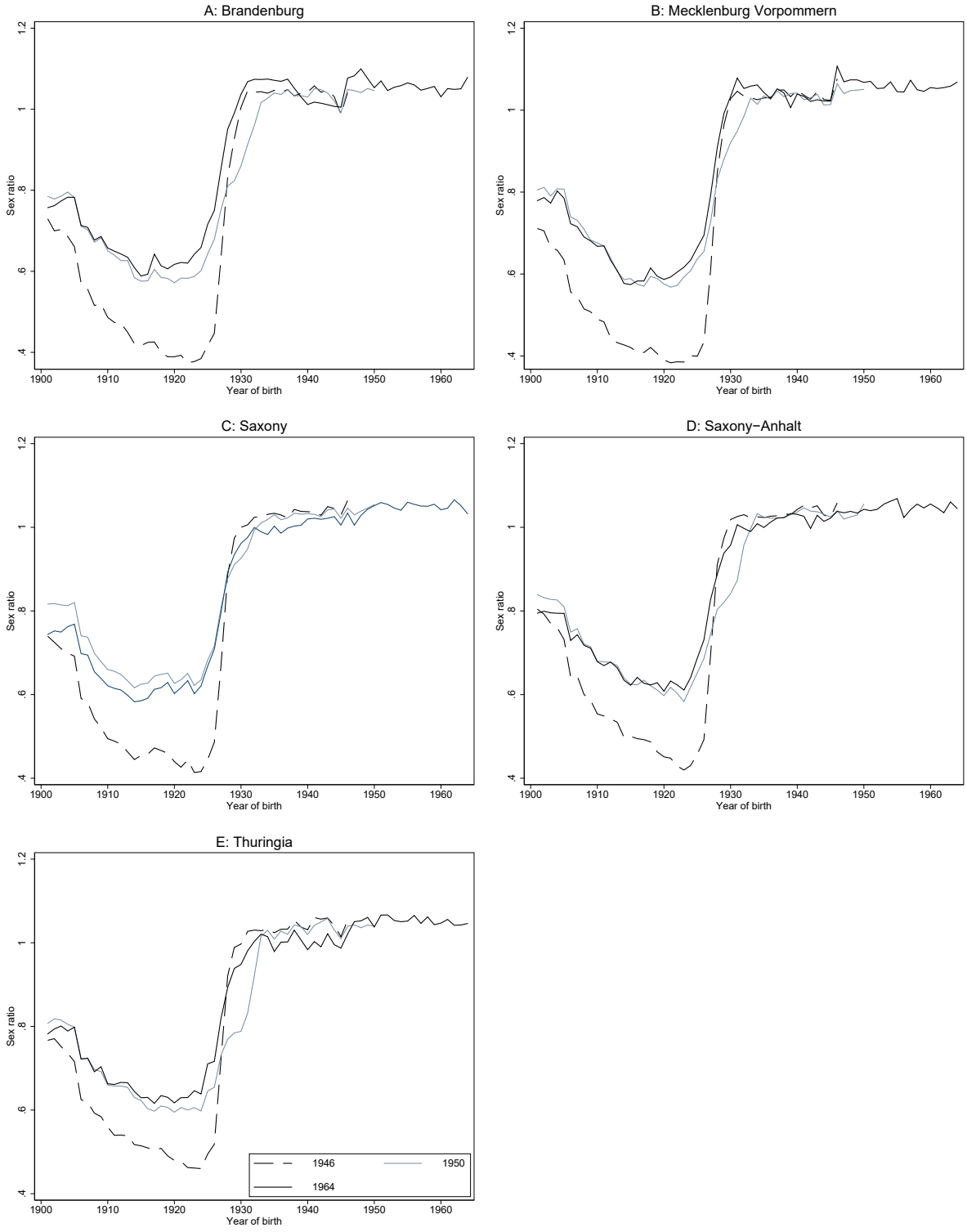
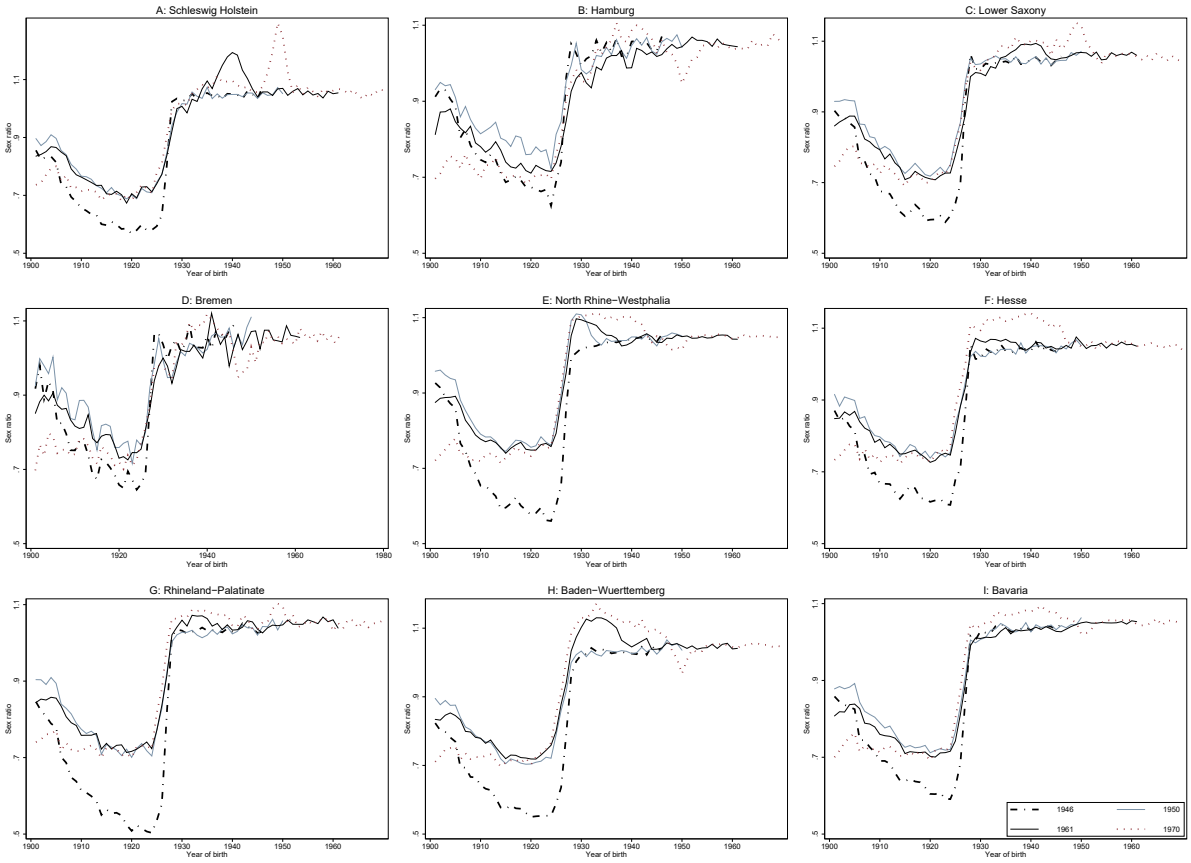


Figure 3.3: Sex ratios in 1946, 1950, 1961 and 1970 by state



3.3.3 Mapping of county-level data into voting-district data

As these data are available at the level of the county, we constructed a mapping of the 1946 counties into the 1990 voting districts. For each voting district individually we searched for its geographical composition and linked it to one or more of the 1990 counties. Those were in turn matched to the 1946 counties.

Similarly, we recoded the 1997 INKAR employment numbers to the 1990 voting district level. Because there is no clear matching of 1990 voting districts to 1997 counties for Saxony-Anhalt, we drop its 13 voting districts for the female labor force participation variable.

3.3.4 Summary statistics

Summary statistics for all variables are reported in Table 3.1. It is clear that women were underrepresented in the 1990 German elections. Out of 2304 first vote candidates,

420 were women (18%) and the share of districts with a female winner was even lower at 12%. While the average number of candidates was a little over 7 per voting district, there were 72 districts with no female candidates. In only 2 districts more than half of the candidates were female. Many incumbent parliamentarians decide to run for office again. In three quarters of the districts, an incumbent is seeking re-election. Female labor force participation, measured as the percentage of women employed in the female population aged 15 to 65, is on average 43.3%. In East-Germany, the participation rate is 5.5 percentage points higher than in West-Germany (47.7% and 42.2% respectively).

Further details of the data construction are available, together with the data, from the authors.

Table 3.1: Descriptive statistics: district and candidate level

Variable		Obs	Mean	Std. Dev.	Min	Max
District level						
Share women	% female nominees	323	.181	.138	0	.667
Sex ratio	Men/Women	323	.795	.047	.638	.909
Catholics ratio	Catholics/Population	323	.38	.292	.045	.98
Vote share left	% votes SPD/PDS/Grüne/B90	323	.43	.092	.235	.659
Incumbency	Dummy, 1 = incumbent present	323	.743	.438	0	1
Female LFP	(women employed/women 15-65)*100	310	43.299	4.249	32.1	54.6
Candidate level						
Female	Dummy, 1 = female	2304	.182	.386	0	1
Age		2304	45.227	11.999	18	86
FV elected	Dummy, 1 = elected in first vote	2304	.14	.347	0	1

3.4 Historical background and preliminary data analysis

In this section, we discuss the main premises on which our analysis rests. We present both historical facts and analyze supporting data. We describe the sources of variation in post-war sex ratios across counties and argue that it is not related to other determinants of later participation of females in the market for politicians. We first show that before the war, sex ratios did not display a distinct geographical pattern at the county level in

Germany. We then describe how the onset of the war, and war itself, changed these sex ratios.

Figure 3.1a, above, shows sex ratios at the county level in May 1939, four months before the war started.[9] There was considerable geographic variation, but there were no distinct patterns – in particular, there are no differences between what would later become East and West Germany. The average sex ratio was about 0.95 with a standard deviation of 0.06. Men who were drafted for mandatory labor or military service at the time of the census were not included in the census. In May 1939, these men were predominantly from birth cohorts 1915–1920 (Statistisches Reichsamt, 1940, pp. 313-314), as can be seen from Figure 3.4.

The figure shows, separately for the later federal states, the sex ratios in May 1939 by year of birth.[10] Note that states are color-coded according to the four post-war occupation zones.[11] There is no systematic pattern across the four zones. Over the course of the war, 17.3 Million men were drafted into the German armed forces (Overmans, 2004, p. 226). In particular, the birth cohorts 1910–1927 were fully drafted (Kroener et al., 1988, p. 727).

In WWII, an estimated 5.3 million German soldiers died (Overmans, 2004, p. 232); around 2.7 million lost their lives at the Eastern front and about 340,000 at the Western front (Overmans, 2004, p. 265).[12] Figure 3.5, which is based on the Census of 1946, illustrates that the war deaths among the drafted cohorts resulted in dramatic declines, and also variation, in sex ratios after the war.

Given the different death rates at the Western and Eastern fronts, it is important to note that where soldiers lived before the war did not matter for where they fought during the war, and divisions were often re-grouped and deployed to different fronts over the course of the war (see e.g. Overmans (2004); Rass and Rohrkamp (2007)). Nevertheless, at the level of individual counties of origin, death rates varied[13] but there is no indication

[9] As discussed above, we could match most counties to the map, but not all; the missing ones are shown in grey.

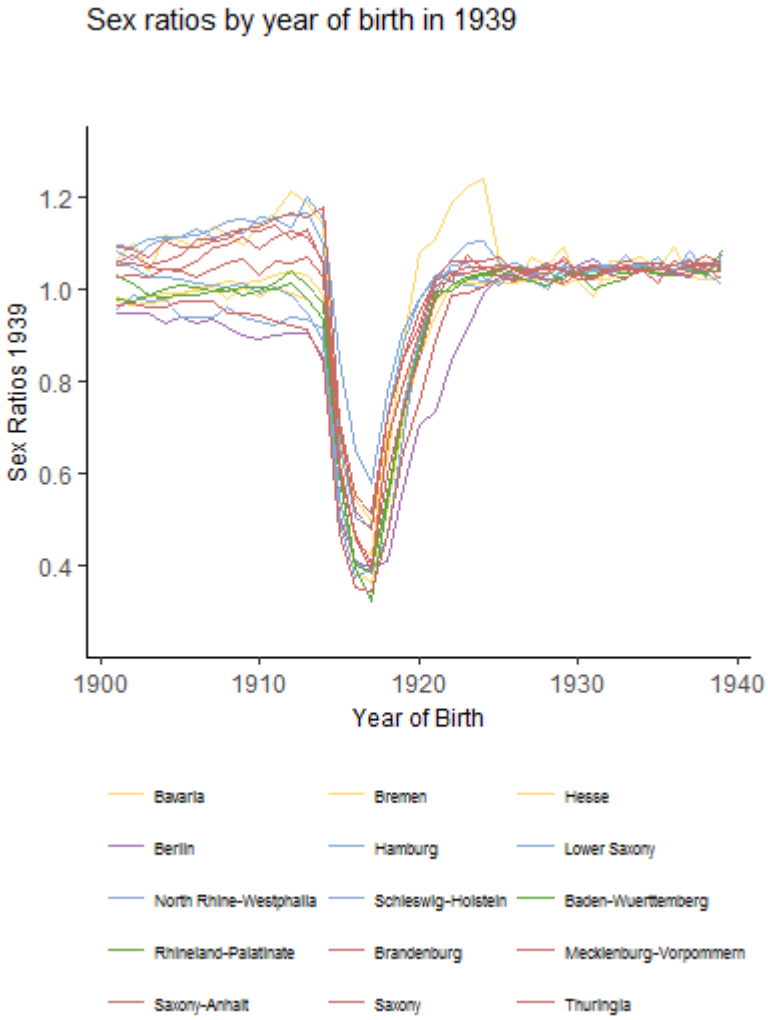
[10] Unfortunately, numbers of men and women by year of birth are available only at the federal state level, so we cannot exploit this variation in our analysis below.

[11] States within the Soviet zone are colored red; states which were part of the American zone are yellow, French zone states are indicated green and the states of the British zone are blue; Berlin is purple.

[12] 104,000 German soldiers died in the Balkans, 151,000 in Italy, 459,000 in war imprisonment, 1.23 Million in the final battles, and 291,000 in other war zones. The number of deaths in war imprisonment entails only deaths registered as such. In these numbers, soldiers are assigned to the place from which the last sign of life was obtained (Overmans, 2004, p. 265).

[13] Data on war-related deaths at this geographic level is not readily available. The *Deutsche Dienststelle*

Figure 3.4: Sex ratios in 1939 by post-unification state and year of birth

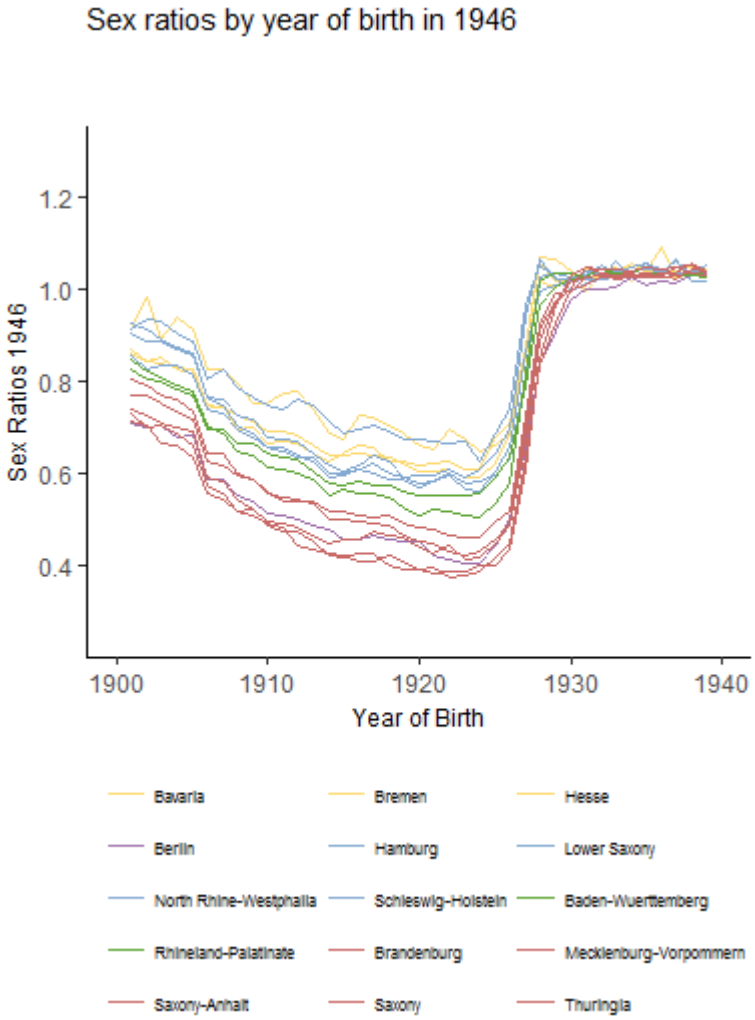


of any underlying systematic cause. In the census of 1946, which was carried out by the four allied powers, the average sex ratio at the county level had dropped to 0.79. For the group of counties where we could obtain a match from 1939 to 1946, the average decrease in the sex ratio was -0.16, with a standard deviation of 0.056.[14] A regression of 1946 sex ratios on 1939 sex ratios using all 705 matched counties yields a significant but small coefficient of 0.30 ($p < 0.00$) and relatively low explanatory power, $R^2 = 0.177$. Thus, less

(WASt) has information on 3.3 Million registered German casualties which can be accessed via individual names. They are, however, not sorted by regions. For a full coverage of military deaths at the county level, all registry offices in Germany would need to be contacted since casualties were registered at the last place of residence before soldiers were drafted.

[14] Altogether, there were 712 counties in the census of 1946 plus the 20 districts of Berlin (Ausschuss der deutschen Statistiker für die Volks– und Berufszählung 1946, 1951, p. 8). Saarland is not contained, because it was not part of four occupational zones in 1946. There are 759 counties within the borders of Germany today in our data from 1939 (no data on the district level for Berlin is available for 1939 in the census results). We were able to get a mapping for 705 counties from 1939 and 1946.

Figure 3.5: Sex ratios in 1946 by post-unification state and year of birth



than a fifth of the cross-county variation in post-war sex ratios is explained by pre-war variation.

The effect of the war on sex ratios was considerably worse in the Eastern part of Germany than in the West, as can be seen in Figure 3.1b, above. As there were no such differences before the war, this pattern is mainly a consequence of front lines at the end of war.[15] When the German army was pushed back in the last months of the war, fights were a lot fiercer on the Eastern front than on other fronts. Towards the end of the war, even old men and young boys from the local population were forced to fight on the Eastern front, with only very basic military training and without appropriate equipment and weapons. This led to very high numbers of casualties in the East (Steinberg, 1991,

[15] See also discussion in Kesternich et al. (2018).

p. 144).

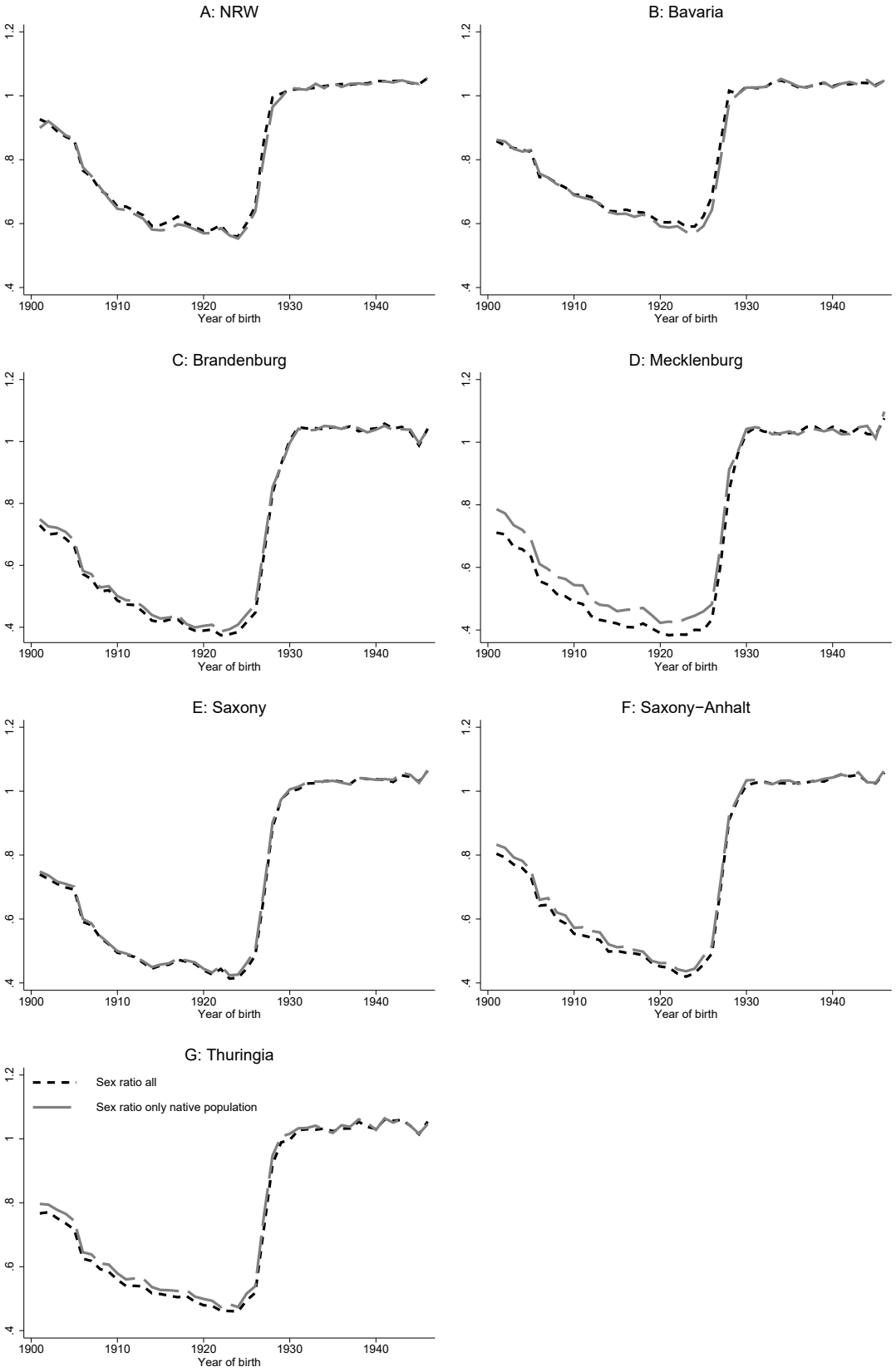
There was also a very strong fear of going into Soviet war imprisonment among German soldiers. During the final weeks of the war, German troops tried to surrender to the Americans rather than the Soviets forces (Henke, 1996). The U.S. Forces estimated that around 3 million German soldiers would surrender to them, while in fact more than 5 million did (Bischof & Ambrose, 1992; Steinberg, 1991, p. 5).

Another consequence of the war was the flow of expellees towards the remaining parts of Germany. In 1946, there were around 9.6 million Germans or ethnic Germans who had not lived in the four occupational zones before the war (Myers & Mauldin, 1952, p. 7). Due to the massive housing shortage after WWII and high inflow of expellees within a short period of time, expellees were primarily distributed according to housing and food availability right after the war (Myers and Mauldin, 1952, p. 41; Kift, 2008, p. 121). Figure 3.6 shows sex ratios for the native population compared to the sex ratios of the overall population at the state level, with and without expellees. Importantly, the sex ratios of the expellees did not have a sizeable impact on sex ratios at the state level.

Moving was heavily restricted for the civilian population in Germany right after the war. Free movement within and across zones was forbidden until 1947 through ordinances by the military government (Hottes & Teubert, 1977, p. 118). Freedom of movement was however not fully restored afterwards. Firstly, the so-called Law 18 (enacted by the Allied Control Council) allowed changing of residence only if housing space could be proved (Hottes & Teubert, 1977, p. 118). Secondly, German cities were heavily destroyed by aerial bombing. Hence, many cities and industrial regions imposed moving bans (*Zuzugssperren*) which were in place until the late 1940s (Krauss, 1985, pp. 149-150; Kift, 2008, p. 121). Freedom of movement was eventually fully restored in 1950 in West Germany.[16]

[16]The legal text of Law 18 as well as exact dates of its repeal can be found in Verfassungen (2004).

Figure 3.6: Sex ratios in 1946 of overall population compared to sex ratios of native population by post-unification state and year



3.5 Results

Our main results concern the effect of post-war (1946) sex ratios on female participation in the market for politicians, as observed in the 1990 elections. We first present regressions that predict the share of female politicians in a voting district using historical sex ratios. We then present separate regressions for whether the elected and non-elected candidates were female. This allows us to differentiate between two channels through which sex ratios can affect female participation – voter demand and candidate supply.

3.5.1 The sex ratio effect on female participation

We first report the results of OLS regressions of the share of female candidates by voting district (in 1990) on the historical sex ratio (in 1946); see Table 3.2. Our baseline regression equation is specified as follows:

$$\%\text{female}_i = \alpha + \beta \ \text{sexratio}_i + X_i'\gamma + \text{state}_i'\delta + \varepsilon_i \tag{3.1}$$

The dependent variable is the share of female nominees in voting district i. The parameter of interest is β. It captures the effect of the historical sex ratios on the supply of female candidates in the 1990 elections. Control variables at the voting district level are contained in the vector X and include a dummy variable for whether the incumbent politician runs again, the share of votes for left wing parties, and the share of Catholics.[17] [18] In addition, we include a vector of state dummies, denoted by state_i. ε_i is the error term. Since we include state dummies in our analysis, the variation in the sex ratio we exploit stems from deviations from the mean sex ratio of a state.

The rationale for selecting the control variables follows arguments from the political science literature. Incumbents have a higher probability of getting re-elected (Hogan, 2004) and are less likely to be female (Matland & Studlar, 1996). As party officials who are skeptical of female participation might decide to nominate female candidates in those

[17] We obtained data on the number of Catholics relative to Protestants from the census of 1946 (Ausschuss der deutschen Statistiker für die Volks– und Berufszählung 1946, 1949b).

[18] We also ran the regressions including the share of expellees as an additional control variable. We do not have information on the number of expellees for the French occupation zone so including this variable reduces our number of observations by 61 voting districts. Since an F-test rejects the hypothesis that adding the share of expellees significantly changes the results, we leave this variable out.

constituencies where they expect them to have the lowest probability of winning, we control for incumbency. Furthermore, gender roles in Catholic regions and conservative parties tend to lead to lower shares of female politicians; see Bieber (2013). We include the share of votes for the three left-wing parties as an additional control variable. By the 1990 elections, voluntary party quotas had been introduced by the left-wing parties SPD and PDS and the Green Party. These quotas were not enforceable and none of the parties reached their quota in the first vote. However, the existence of such quota could lead to a correlation of the share of left wing votes with both the share of female candidates and the sex ratio.

Results for the regressions that include all constituencies are presented in Panel 3.2a. The first column shows the baseline specification, a regression of the share of female nominees on the 1946 sex ratio. Control variables are added sequentially in columns 2 to 4. The results are robust to the addition of control variables; in fact they strengthen the sex ratio effect. When we consider the specification that includes all control variables, the effect of the sex ratio on the share of female nominees is significant at the 5% level.[19] A decrease of the sex ratio by one standard deviation (i.e., by 0.047) is associated with an increase in the share of women nominated by approximately 3.9 percentage points. To illustrate, a decrease in the sex ratio from 0.82 to 0.72, which are the values of North-Rhine Westphalia and Brandenburg, respectively, is associated with an increase in the share of female nominees by 8.3 percentage points. As female representation in the Bundestag after the 1990 elections was around 20 percent and the share of female nominees for the first vote was around 18 percent, the effect of historic sex ratio imbalances is of relevant size.[20]

Panel 3.2b reports results for specifications that drop the districts located in the State of Bavaria. Bavaria has a special role in the German political system. First, it has a very high share of Catholics, and second, the conservative party (CSU) has had a commanding share of the vote over much of the post-war period. Leaving out the 45

[19] We also ran an OLS regression that does not include the share of left votes as a control variable. We find that our results are robust to leaving out this variable. The sex ratio effect remains statistically significant and of similar magnitude as in the base specification reported here.

[20] Separate regressions for East and West Germany suggest that the sex ratio effect is larger in West Germany (see Table 3.10). The estimated West German sex ratio effect is -0.95 and we find an insignificant sex ratio effect in East Germany. However, since there are only 80 voting districts in East Germany, its regression has a small sample size.

Table 3.2: OLS regression of the share of female nominees in a constituency on 1946 sex ratios

(a) Including Bavaria

	(1) % female	(2) % female	(3) % female	(4) % female
Sex ratio	-0.609*	-0.609*	-0.741**	-0.831**
	(0.329)	(0.329)	(0.356)	(0.360)
Incumbency		-0.001	-0.000	0.001
		(0.019)	(0.019)	(0.019)
Vote share left			0.174	0.238*
			(0.124)	(0.133)
Catholics ratio				0.051
				(0.054)
Constant	0.666**	0.666**	0.705***	0.729***
	(0.257)	(0.257)	(0.266)	(0.266)
Observations	323	323	323	323
R-squared	0.036	0.036	0.043	0.046
State dummies	YES	YES	YES	YES

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

(b) Excluding Bavaria

	(1) % female	(2) % female	(3) % female	(4) % female
Sex ratio	-0.864**	-0.864**	-0.942**	-1.006***
	(0.343)	(0.345)	(0.368)	(0.373)
Incumbency		-0.005	-0.005	-0.005
		(0.020)	(0.020)	(0.020)
Vote share left			0.101	0.159
			(0.119)	(0.133)
Catholics ratio				0.052
				(0.057)
Constant	0.867***	0.870***	0.894***	0.899***
	(0.268)	(0.269)	(0.276)	(0.277)
Observations	278	278	278	278
R-squared	0.045	0.045	0.047	0.051
State dummies	YES	YES	YES	YES

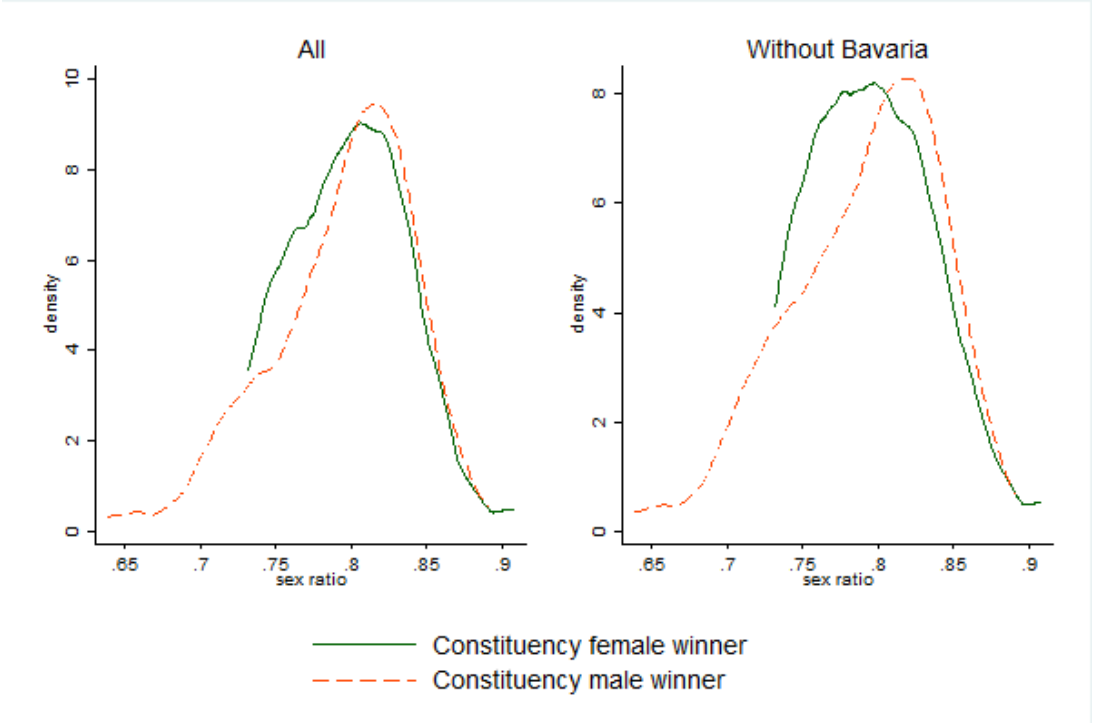
Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Bavarian constituencies from the sample increases the sex ratio effect. For each 0.01 decrease in 1946 sex ratio, the share of female nominees now increases by 1 percentage point.

3.5.2 Candidate supply vs. voter demand effects

We now investigate whether female candidates were more likely to win in districts that had lower sex ratios. Figure 3.7 shows 1946 sex ratio densities, separately for voting districts with female and male winners in 1990. Sex ratios are substantially smaller than 1 for each constituency, illustrating the devastating effects of WWII on the German male population. One can see in the graph that constituencies where sex ratios were lower, i.e., where there were relatively more women than men in 1946, a female candidate was more likely to receive the majority of votes. This is especially the case when we exclude the 45 Bavarian constituencies. The null hypothesis of identical distributions, however, cannot be rejected (using non-parametric Kolmogorov-Smirnov tests).

Figure 3.7: Sex ratio densities for constituencies with female/male winners



To test whether elected candidates were more likely to be female in constituencies with lower historical sex ratios, we slightly modify regression equation (3.1) and estimate

a probit regression,

$$\text{female}_j = \Phi(\alpha + \beta\ \text{sexratio}_i + \lambda\ \text{age}_j + X_i'\gamma + \text{state}_i'\delta + \varepsilon_j)\,. \tag{3.2}$$

In this specification, the outcome is a binary variable that equals 1 if candidate j is female. Note that this analysis is conducted at the candidate rather than the voting district level. We can therefore add the candidate's age as an additional control variable. The two voting district control variables (vote share left and catholics ratio) remain the same as in regression equation (3.1). However, the "incumbent opposer" dummy differs slightly from the incumbency dummy. While the incumbency dummy is defined at the voting district level and indicates the presence of an incumbent politician in the constituency, the incumbent opposer dummy is defined at the candidate level. It indicates whether one of the other candidates in the constituency is an incumbent politician.[21, 22]

Table 3.3 presents the results of probit regressions for two subsets of the sample, all candidates and candidates elected by first vote. Columns 1 and 2 show the results for regressions using data on all nominees. We find that the sex ratio effect is negative and significant at the 5% level for the whole sample and significant at the 1% level for candidates outside of Bavaria. Average marginal effects of the sex ratio are displayed in Panel 3.3b and are in line with the results in section 3.5.1. A one percentage point increase in the 1946 ratio of men relative to women is associated with a 0.833 percentage point decrease in the likelihood that a nominee is female, a sex ratio effect similar in interpretation and size as the results from Table 3.2.

Note that we do not find a significant effect of the sex ratio on the likelihood of a woman winning the election in first vote, irrespective of whether Bavarian constituencies are included in the estimation (columns 3 and 4).[23] This suggests that while women

[21] To illustrate the difference, suppose that candidate A in voting district B is the (only) incumbent. The incumbency dummy will be equal to 1 for voting district B since there is an incumbent candidate. The incumbent opposer dummy variable, on the other hand, is equal to 0 for candidate A since none of the other candidates in voting district B are incumbents.

[22] A linear probability model with the same outcome variable and regressors also predicts a negative relationship between sex ratios and the probability that a nominated candidate is female. Similar to the results from equation (3.1), a 0.01 increase in the sex ratio increases the probability of a female nomination by 0.84 percentage points ($p < 0.05$). The linear probability model finds no significant relationship between the district-level sex ratio and the gender of an elected candidate.

[23] There are 72 constituencies with only male nominees. When we re-run the probit regressions for the last two columns of Table 3 without those 72 constituencies where women cannot be elected, results remain far from being significant.

Table 3.3: Probit regression of the probability that candidate is female on 1946 sex ratio. Columns 1-2: candidate nominated for first vote, errors are clustered at the voting district level. Columns 3-4: candidate elected by first vote.

(a) Coefficient estimates

	Nominated		Elected	
	(1)	(2)	(3)	(4)
	Incl Bavaria	Excl. Bavaria	Incl Bavaria	Excl. Bavaria
Sex ratio	-3.176**	-4.006***	-0.262	-1.117
	(1.278)	(1.310)	(4.761)	(4.930)
Age	0.001	-0.001	-0.037***	-0.039***
	(0.002)	(0.003)	(0.013)	(0.014)
Incumbent opposer	0.121*	0.088	0.421	-0.022
	(0.066)	(0.070)	(0.355)	(0.361)
Vote share left	0.894*	0.556	0.968	-0.089
	(0.467)	(0.473)	(1.647)	(1.601)
Catholics ratio	0.197	0.191	-0.442	-0.505
	(0.198)	(0.209)	(0.605)	(0.640)
Constant	1.097	1.966*	0.735	1.952
	(0.956)	(0.990)	(3.427)	(3.615)
Observations	2,304	1,960	323	278
State dummies	YES	YES	YES	YES

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

(b) Average marginal effects

	Nominated		Elected	
	(1)	(2)	(3)	(4)
	Incl Bavaria	Excl. Bavaria	Incl Bavaria	Excl. Bavaria
Sex ratio	-0.833**	-1.069***	-0.048	-0.205
	(0.334)	(0.346)	(0.868)	(0.903)
Observations	2,304	1,960	323	278

Standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

were more likely to run for a seat in parliament in constituencies with lower historical sex ratios, voters were not more inclined to vote for them. As discussed above, the fact that there are few female candidates can be due to women's preferences, voters' preferences, or the preferences of parties leaders, see e.g. Casas-Arce and Saiz (2015). These three

explanations are difficult to disentangle, but our results provide some indication that party leaders' preferences are not decisive. As mentioned above, we would expect more women to be elected if party leaders were holding back women at the expense of votes before. Since this is not the case, we argue that party leaders' preferences are unlikely to be the driving factor behind our findings.

3.6 Channels for long-lasting sex-ratio effects

There are two main channels through which lower sex ratios after WWII could have led to an increase in the share of female politicians many years later: One is the increase in female employment in regions where many men were missing after WWII and the other an increase in the share of female voters in these regions.

We first discuss the effects of WWII on female labor force participation. During the war, the number of women who worked had increased only slightly – from 14.626 million in May 1939 to 14.897 million in September 1944 (i.e., less than 2%) (Kaldor, 1945). The fact that female employment increased so little during the war can be explained, in part, by the Nazi government's heavy reliance on forced labor and prisoners of war (Kaldor, 1945). This changed of course after the war when the Allied forces freed the surviving forced laborers. A second reason is that under the Nazi ideology, children and housework were seen as the essential activities of women (Moeller, 1993).

At the end of the war, there was a severe labor shortage, with so many men still being away (Niehuss, 2001, p. 77). Whether a woman looked for work in the official labor market in the immediate post-war period depended to a large extent on the individual situation and the availability of other options to secure survival of the family. Many women worked in the informal labor market. With wages frozen at the pre-war level of 1939, a flourishing black market emerged (Moeller, 1993, pp. 22-23; Niehuss, 2001, p. 63). One reason for working in the formal sector was the desire to obtain a better ranking in the food rationing system, that is, to get out of category V which comprised housewives, pensioners, etc. and was called the "starvation or ascension category" (Moeller, 1993, p. 22). For divorced women (who only received alimony under certain conditions) or widows, employment was very important as well (Niehuss, 1993, p. 323; Niehuss, 2001, pp. 63-64).

Overall, around 500,000 more women were in the labor force in October 1946 than in 1939 (Ausschuss der deutschen Statistiker für die Volks– und Berufszählung 1946, 1953, p. 19).[24] At the state level, a higher surplus of women in the population went along with a higher female employment rate in the census of 1946 (Ausschuss der deutschen Statistiker für die Volks– und Berufszählung 1946, 1953, p. 16).

Many women had to take over the position of the head of household. As the food situation worsened drastically towards the end of the war and remained at an extremely poor level in the immediate post-war years, women had to bear the burden of ensuring the survival of their families. In the 1950 Census, 1.7 million of the overall 15.4 million households in West Germany consisted of women who lived alone and in 2.1 million households women were the head of household (Moeller, 1993, p. 32).

The phenomenon of women taking the economic responsibility for themselves and their families was called "forced emancipation" (Moeller, 1993, pp. 29 and 32). Society in Western Germany accepted this phenomenon, but only as it was seen as a response to the extraordinary hardships caused by the highly unbalanced sex ratios (the German term coined at the time was *Frauenüberschuss*, or "excess women") (Moeller, 1993, p. 28). In general, however, the West German society remained quite conservative in their views on female roles (Moeller, 1993). At the other extreme, in Eastern Germany, the government tried to mobilize the female labor force as much as possible. In 1970, around 49% of women aged 15 to 60 were working in West Germany. In East Germany, around 69% of women in the same age range were economically active in 1971 (Minnesota Population Center, 2017).[25]

We illustrate the relevance of this channel by regressing female labor force participation (LFP) in 1997, the year closest to the election data for which we have county-level information on female employment indicators, on post-war sex ratios. Information on female LFP in earlier years is only available at the state-level. Female LFP is measured as the share of employed women in the female population aged 15 to 65. The regression results in Table 3.4 show a negative and significant correlation between post-war sex ratios

[24] This increase of 500,000 women in the work force compared to 1939 results from: (+) 65,205 workers, (+) 689,821 civil servants and employees, (-) 661,883 family workers, (+) 405,446 self-employed. This strong increase in self-employed women stems from the fact that women had to take over the family business temporarily or permanently if their husband was still in war imprisonment or had died in war (Ausschuss der deutschen Statistiker für die Volks– und Berufszählung 1946, 1953, p. 28).

[25] The underlying census data is provided by the Federal Statistical Office, Germany.

Table 3.4: OLS regression of 1946 sex ratio effect on 1997 female labor force participation.

	(1) Female LFP	(2) Female LFP	(3) Female LFP
Sex ratio	-0.538***	-0.376***	-0.411***
	(0.0512)	(0.0777)	(0.0689)
East		2.385***	
		(0.674)	
Constant	86.15***	72.70***	77.05***
	(4.072)	(6.282)	(5.435)
Observations	310	310	287
R-squared	0.353	0.374	0.664
State dummies	NO	NO	YES

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

and female LFP and suggest that a 1 percentage point increase in 1946 sex ratios led to a decrease of 0.538 number of employed women per 100 women aged 15 to 65. The sex ratio effect is weaker but still significant at the 1% level after controlling for East-West or state differences.

The effect of unbalanced sex ratios on female nominees is larger than that on female LFP, in both absolute and relative terms. A one percentage point decrease of the sex ratio corresponds to a 0.609 percentage point increase in female nominees and a 0.411 percentage point increase in female LFP in a specification where we only control for state effects. Since female labor force participation has a higher base rate, this translates into a relative increase of 3.36% in the share of female nominees and a 0.95% increase in female labor force participation (relative to the mean).

Parties in regions with low sex ratios might have adopted more female-friendly policies, and thus would have been more open to female politicians, also for a second reason: Low sex ratios translate into higher shares of female voters. A quote by Elisabeth Selbert, one of the "mothers of the German constitution", illustrates that political parties were well aware of this fact: A "woman, who during the war years stood atop of the rubble and replaced men at the workplace, [and] has a moral right to be valued like a man." And this especially given the highly unbalanced sex ratios which "left 170 women voters for every 100 men" (Moeller, 1993, p. 51).

We believe that both these channels might have contributed to the sex-ratio effects we reported in the previous section. Distinguishing between these channels is, unfortunately, difficult, as we are not aware of any data that would allow us the assess the relevance of the second channel.

3.7 Robustness checks

One concern with our analysis so far is that the 1990 elections were special as they were conducted just after unification. As a robustness check, we therefore re-estimate the regressions reported above with the dependent variables replaced with the outcomes of the 1994 election. The results of OLS regressions of the share of female nominees in a constituency in the 1994 Bundestag elections on 1946 sex ratios are presented in Table 3.5 (all states) and Table 3.6 (excluding Bavaria). Table 3.7 reports the result of Probit regressions for the probability that a candidate is female in the 1994 Bundestag elections on 1946 sex ratio.[26] The effects of the post-war sex ratios on the 1994 election outcomes in all these robustness regressions are quantitatively very similar to those we obtained for the 1990 elections.

In the main part of the analysis, we did not use data on the elections before 1990 because the political system was not comparable in East and West Germany whereas from 1990 on, data from East and West can be combined. This is important since our sample is relatively small, and also because we have much more variation in the sex ratios when we also include East Germany. Before the public discussion about the low shares of female politicians and the introduction of quotas picked up in the 1980s, base rates were also small. The voting system in the first election after WWII, held in 1949, was not comparable to the current one. In 1953, only 9.7 percent of all candidates were female and only 3.7 percent of constituencies were won by women (Der Bundeswahlleiter, 2017b). The share of constituencies won by women stayed quite low throughout the 60s and 70s. It exceeded the 5 percent level not until 1987 when the share was 7.3 percent (Der Bundeswahlleiter, 2017b). Changes in female voting behavior and increased participation in political activities were associated with the introduction of party quotas in the 1980s in

[26] In these probit regressions, three states are dropped because they had only male politicians; this reduces the sample size by 39 voting districts.

Table 3.5: OLS regression of the share of female nominees in a constituency in the 1994 Bundestag elections on 1946 sex ratios. Sample includes Bavarian constituencies.

	(1) % female	(2) % female	(3) % female	(4) % female
Sex ratio	-0.648*	-0.640*	-0.720*	-0.697*
	(0.383)	(0.386)	(0.404)	(0.422)
Incumbency		0.057***	0.056**	0.056**
		(0.022)	(0.022)	(0.022)
Vote share left			0.128	0.110
			(0.135)	(0.150)
Catholics ratio				-0.014
				(0.054)
Constant	0.747**	0.695**	0.705**	0.701**
	(0.304)	(0.305)	(0.308)	(0.311)
Observations	323	323	323	323
R-squared	0.030	0.050	0.053	0.053
State dummies	YES	YES	YES	YES

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Table 3.6: OLS regression of the share of female nominees in a constituency in the 1994 Bundestag elections on 1946 sex ratios. Sample excludes Bavarian constituencies.

	(1) % female	(2) % female	(3) % female	(4) % female
Sex ratio	-0.747*	-0.737*	-0.793*	-0.786*
	(0.404)	(0.408)	(0.428)	(0.442)
Incumbency		0.050**	0.049**	0.049**
		(0.023)	(0.023)	(0.023)
Vote share left			0.084	0.077
			(0.144)	(0.160)
Catholics ratio				-0.006
				(0.056)
Constant	0.825**	0.777**	0.787**	0.787**
	(0.321)	(0.322)	(0.326)	(0.326)
Observations	278	278	278	278
R-squared	0.035	0.052	0.053	0.053
State dummies	YES	YES	YES	YES

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Table 3.7: Probit regression of the probability that candidate is female in the 1994 Bundestag elections on 1946 sex ratio. Columns 1-2: candidate nominated for first vote, errors are clustered at the voting district level. Columns 3-4: candidate elected by first vote.

(a) Coefficient estimates

	Nominated		Elected	
	(1)	(2)	(3)	(4)
	Incl Bavaria	Excl. Bavaria	Incl Bavaria	Excl. Bavaria
Sex ratio	-2.483*	-2.823**	-3.790	-5.823
	(1.308)	(1.352)	(4.621)	(4.879)
Age	0.001	0.000	-0.025**	-0.028**
	(0.003)	(0.003)	(0.012)	(0.013)
Incumbent opposer	0.319***	0.282***	-0.234	-0.103
	(0.066)	(0.070)	(0.505)	(0.531)
Vote share left	0.316	0.183	3.179*	2.722
	(0.464)	(0.492)	(1.732)	(1.722)
Catholics ratio	-0.037	0.002	-0.544	-0.660
	(0.168)	(0.174)	(0.583)	(0.636)
Constant	0.852	1.222	1.921	3.873
	(0.981)	(1.019)	(3.343)	(3.582)
Observations	2,265	1,934	284	239
State dummies	YES	YES	YES	YES

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

(b) Average marginal effects

	Nominated		Elected	
	(1)	(2)	(3)	(4)
	Incl Bavaria	Excl. Bavaria	Incl Bavaria	Excl. Bavaria
Sex ratio	-0.751*	-0.860**	-0.814	-1.287
	(0.395)	(0.411)	(0.989)	(1.071)
Observations	2,265	1,934	284	239

Standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

the Green Party as well as SPD (Kolinsky, 1991). Kolinsky (1991) summarizes the spirit of the time in the following way:

"The introduction of women's quotas has been preceded by a varied history of controversy and debate about the role of women in the respective political parties. In a broader context, however, quotas were designed to meet the expectations about equal opportunities

among West German women of the post-war generations, that equal educational and professional qualifications should be matched by equal chances to participate in society and politics." (Kolinsky, 1991, p. 56)

As a further robustness check, we obtained data for the 1987 elections in West Germany (and also on incumbents elected in 1983). The estimation of a sex ratio effect in this election corroborates our expectations as shown in Table 3.8.

Table 3.8: OLS regression of the share of female nominees in a constituency in the 1987 Bundestag elections on 1946 sex ratios.

	(1)	(2)
	% female	% female
Sex ratio	-0.194	-0.443
	(0.523)	(0.568)
Incumbency	-0.008	-0.014
	(0.029)	(0.034)
Vote share left	0.150	0.235
	(0.199)	(0.226)
Catholics ratio	-0.042	0.001
	(0.067)	(0.075)
Constant	0.334	0.481
	(0.377)	(0.410)
Observations	243	198
R-squared	0.020	0.015
state dummies	YES	YES
incl. Bavaria	YES	NO

Standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

The sex ratio effect shows the same negative sign as in the 1990 elections and is of smaller magnitude. However, we obtain larger standard errors and the effect is not significantly different from zero. This is at least partially a consequence of the decrease in sample size by 80 voting districts. Hence, a tentative explanation is that the public debate about female representation in political offices of the 1980s combined with the longer-run social norms effects had an impact on the outcome variable we consider that is large enough to be measured. Thus, there might well have been similar effects even before 1990, but these were likely smaller and less precisely estimated because of the smaller sample.

To shed some light on the effect of quotas, we split up the candidates between right (no quotas) and left parties (quotas) (see Table 3.9). We find that the sex ratio effect for nominees is significant for the SPD/PDS/Green party while there is no significant effect for CDU/CSU/FDP candidates.

Table 3.9: Probit regressions of the probability that candidate, affiliated with a right (CDU/CSU/FDP) or left (SPD/PDS/Grüne) wing party, is female on 1946 sex ratio.

	(1) CDU/CSU/FDP	(2) SPD/PDS/Grüne
Sex ratio	0.101	-4.495*
	(3.015)	(2.517)
Age	-0.020***	-0.007
	(0.007)	(0.006)
Incumbent opposer	0.276**	0.108
	(0.131)	(0.118)
Vote share left	0.085	0.275
	(1.021)	(0.848)
Catholics ratio	0.087	0.370
	(0.383)	(0.331)
Constant	-0.154	2.962
	(2.253)	(1.874)
Observations	646	700
state dummies	YES	YES

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

As sex ratios were particularly low in East Germany, there is a concern that our findings might be driven by East Germany alone. This effect might be exacerbated by the fact that the share of female politicians was higher in East Germany where policies that encouraged female labor market participation had been implemented. We address this concern by presenting separate regressions for East and West Germany in Table 3.10.

The sex ratio effect on the share of female nominees in the 1990 election is negative and statistically significant also when only data from West Germany are used. It is also negative but statistically insignificant for East Germany where the sample is rather small. This being said, we would have expected a smaller effect for East Germany because of gender equality policies that reduced the scope for the channel through which low sex ratios might have encouraged labor-market participation.

Table 3.10: OLS regression of the share of female nominees in a constituency on 1946 sex ratios: West- and East Germany.

	(1) West	(2) East
Sex ratio	-0.950**	-0.650
	(0.429)	(0.745)
Incumbency	0.00324	-0.00541
	(0.0224)	(0.0344)
Vote share left	0.306*	0.132
	(0.184)	(0.191)
Catholics ratio	0.0672	-0.206
	(0.0571)	(0.277)
Constant	0.788***	0.613
	(0.303)	(0.512)
Observations	243	80
R-squared	0.048	0.053
State dummies	YES	YES

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

As a further robustness check, we investigate whether those cohorts that have been affected directly drive our results: In the 1990 elections, there were still some politicians who experienced the war-induced sex ratio imbalances themselves. The last fully-drafted cohort was 1927, so the post-WWII sex ratio imbalances would have had a direct effect on the gender composition of politicians over the age of 63. While the share of female politicians is slightly higher for those cohorts than the average in the 1990 election (21% for politicians born 1927 or before and 18% for politicians born 1928 or later), this difference is not statistically significant and it also does not drive the sex ratio effect. When we re-run the probit regression from column 1 of Table 3.3a for the politicians born in 1928 or later (i.e. excluding those directly affected), we obtain a sex ratio effect of the same order of magnitude and level of significance (see Table 3.11).

Relatedly, we analyze the impact of the sex ratio on the age composition of the electorate. Regions with a lower share of elderly population at the time of the 1990 elections might have been those with many males dead or missing during the war. If such a link between the historic sex ratio and the age composition of the electorate existed, this would be an additional channel for the sex ratio's impact on the share of female nominees since

Table 3.11: Probit regression of the probability that elected candidate is female on 1946 sex ratio. Sample is restricted to candidates born in 1928 or after.

	Female
Sex ratio	-3.570***
	(1.351)
Age	-0.002
	(0.003)
Incumbent opposer	0.122*
	(0.071)
Vote share left	0.894*
	(0.495)
Catholics ratio	0.251
	(0.207)
Constant	1.469
	(1.004)
Observations	2,121
state dummies	YES

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

a younger population might be more prone to accepting female politicians. To check this possibility, we regress the mean age of the population in 1995, the year closest to the election for which we have this information available on the district level, on the 1946 sex ratio. Table 3.12 shows that the effect of the sex ratio on the age composition of the electorate is insignificant.[27]

Table 3.12: OLS regression of the mean population age in 1995 on 1946 sex ratio.

	Mean age pop.
Sex ratio	1.853
	(1.873)
Constant	38.174***
	(1.501)
Observations	302
R-squared	0.003

Standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

[27]In a similar fashion, we checked if constituency size is affected by the 1946 sex ratio and we do not find a significant correlation between these two variables either.

3.8 Conclusions

We exploit random variation in post-WWII sex ratios in Germany to analyze the long-term effects of gender imbalances on female labor force participation. The particular outcome we study is the representation of women in the German market for politicians. We find significant effects of the 1946 sex ratio on the share of female nominees in the 1990 federal election. This result provides evidence for the existence of a mechanism linking past gender imbalances to future participation of women in professional job markets. A plausible mechanism is based on the observation that due to a shortage of male workers after the war, women had to take up jobs that were previously ascribed to men, which in turn might have changed perceptions about the careers that are open to, and should be pursued by, women. Further, the data indicate that voter demand effects were not as strong as candidate supply effects.

Taken together, we interpret our results on female labor market participation, the share of female candidates, and the share of females elected as follows. The effect of historical sex ratios on female labor market participation cannot be driven by those directly affected since they will long have retired by 1997. Thus, we argue that the effect is likely to be driven by a change in norms about female labor market participation. We see that the increase in female participation in the market for politicians is more pronounced than the change in female labor market participation overall. Thus, female politicians profit more than proportionally from changes in norms about female labor market participation. However, while women are more likely to run for office, voters are not significantly more likely to vote for these female candidates.

Exposure to female leaders can change biased perceptions about women's ability as politicians (Beaman et al., 2009). If voters are risk-averse and they were never exposed to females in leading positions, they may refrain from voting for female politicians (Aigner & Cain, 1977). More generally, female labor force participation is in part determined by cultural beliefs about the role of women. These beliefs can be permanently altered through imbalances in sex ratios and the resulting shift in female labor supply (Fernández et al., 2004; Teso, 2019).

While we find the evidence for the existence of a link between past gender imbalances

and current participation of women in a highly competitive labor market compelling, the limited data on intermediate outcomes do not allow us to investigate the specific causal channels that created this association. We believe that this is an important topic for future research on gender imbalances. Nevertheless, our findings highlight the possibility that role models might have an encouraging effect on female participation in professional labor markets even today.

3.A Data appendix

3.A.1 Census data

(a) Census data 1939 (county level)

Statistisches Reichsamt (1941). Volkszählung. Die Bevölkerung des Deutschen Reichs nach den Ergebnissen der Volkszählung 1939. Heft 2: Geschlecht, Alter und Familienstand der Bevölkerung des Deutschen Reichs. Tabellenteil, Statistik des deutschen Reiches, volume 552, issue 2.

(b) Census data 1946 (county level)

Ausschuss der deutschen Statistiker für die Volks- und Berufszählung 1946 (1949a): Volks- und Berufszählung vom 29. Oktober 1946 in den vier Besatzungszonen und Gross-Berlin: Deutsches Gemeindeverzeichnis. Berlin and Munich: Duncker and Humblot.

Ausschuss der deutschen Statistiker für die Volks- und Berufszählung 1946 (1949b): Volks- und Berufszählung vom 29. Oktober 1946 in den vier Besatzungszonen und Gross-Berlin: Volkszählung Tabellenteil. Berlin and Munich: Duncker and Humblot.

Statistisches Amt der Vereinigten Wirtschaftsgebietes (1950). Die Flüchtlinge in Deutschland. Ergebnisse der Sonderauszählung aus der Volks- und Berufszählung vom 29. Oktober 1946, Statistische Berichte, Nr. VIII/0/4, Wiesbaden.

Statistisches Zentralamt für die sowjetische Besatzungszone (1949). Volks - und Berufszählung vom 29. Oktober 1946 in der sowjetischen Besatzungszone. Band III: Landes- und Kreisstatistik, issues 1-5.

(c) References for graphs

(I) Figure 2 (Sex ratios by state - I)

Ausschuss der Deutschen Statistiker für die Volks- und Berufszählung 1946 (1949b). Volks- und Berufszählung vom 29. Oktober 1946 in den vier Besatzungszonen und Gross-Berlin. Volkszählung: Tabellenteil. Berlin and Munich: Duncker & Humblot.

Bundesarchiv Berlin, Referat DDR 3, Archivsignatur DE 2/33229, Die Bevölkerung nach Geburtsjahren für das gesamte Gebiet der DDR, für die Länder (1 x ohne Wismut und Volkspolizei, 1 x mit Wismut und Volkspolizei) sowie für die Bezirke und Kreise (nach dem Gebietsstand 04.12.1952).

Staatliche Zentralverwaltung für Statistik (1967). Ergebnisse der Volks- und Berufszählung am 31. Dezember 1964, Staatsverlag der Deutschen Demokratischen Republik.

(II) Figure 3 (Sex ratios by state - II)

Ausschuss der Deutschen Statistiker für die Volks- und Berufszählung 1946 (1949b). Volks- und Berufszählung vom 29. Oktober 1946 in den vier Besatzungszonen und Gross-Berlin. Volkszählung: Tabellenteil. Berlin and Munich: Duncker & Humblot.

Bayerisches Statistisches Landesamt (1952). Volks- und Berufszählung am 13. September 1950 in Bayern. Volkszählung. I. Teil: Gliederung der Wohnbevölkerung. Beiträge zur Statistik Bayerns, vol. 171, 26-29.

Hessisches Statistisches Landesamt (1952). Die Struktur der Bevölkerung in

Hessen. Endgültige Ergebnisse der Volkszählung vom 13. September 1950. Beiträge zur Statistik Hessen. Sonderreihe Volkszählung 1950, issue 3, 12-15.

Niedersächsisches Amt für Landesplanung und Statistik (1953). Die Bevölkerung am 13. September 1950 nach Altersjahren und Geschlecht. Statistisches Jahrbuch für Niedersachsen, 1952.

Statistisches Landesamt Baden-Württemberg (1959). Ergebnisse der Volks- und Berufszählung vom 13. September 1950. Statistik von Baden-Württemberg, vol. 4.1, 46-51.

Statistisches Landesamt der Freien und Hansestadt Hamburg (1953). Die Volkszählung in Hamburg am 13. September 1950. Statistik des Hamburgischen Staates, issue 36, 13-14.

Statistisches Landesamt Nordrhein-Westfalen Düsseldorf (1952). Die Wohnbevölkerung in Nordrhein-Westfalen nach Alter, Familienstand und Wohnort am 1. September 1939. Ergebnisse der Volkszählung vom 13. September 1950. Beiträge zur Statistik des Landes Nordrhein-Westfalen, Sonderreihe Volkszählung 1950, issue 3, 16-19.

Statistisches Landesamt Rheinland-Pfalz (1953). Volkszählung am 13. September 1950. Die Wohnbevölkerung in Rheinland-Pfalz, Landesergebnisse. Statistik von Rheinland-Pfalz, vol. 17, Bad Ems.

Statistisches Landesamt Bremen (1953). Die Volks- und Berufszählung am 13.9.1950 im Lande Bremen. Statistische Mitteilungen aus Bremen. Sonderheft 4.

Statistisches Landesamt Schleswig-Holstein (1953). Die Bevölkerungsstruktur in Schleswig-Holstein. Ergebnisse der Volkszählung vom 13. September 1950.

Statistik von Schleswig-Holstein, issue 8, 4-5.

Statistisches Bundesamt Wiesbaden (1966). Volks- und Berufszählung vom 6. Juni 1961 Heft 4. Bevölkerung nach Alter und Familienstand, Fachserie A, Bevölkerung und Kultur. W. Kohlhammer GmbH, Stuttgart und Mainz, p. 70-117.

Statistisches Bundesamt (2017). Bevölkerung nach Bundesländern 1970, Wiesbaden.

(III) Figure 4 (Sex ratios in 1939)

Statistisches Reichsamt (1940a). Weitere Teilergebnisse der Volks- und Berufszählung vom 17.Mai 1939. Die ständige Bevölkerung nach Geschlecht, Alter und Familienstand. Wirtschaft und Statistik, vol. 20, issue 12.

Statistisches Reichsamt (1940b). Weitere Teilergebnisse der Volks- und Berufszählung vom 17.Mai 1939. Die ständige Bevölkerung nach Geschlecht, Alter und Familienstand. Wirtschaft und Statistik, vol. 20, issue 15.

Statistisches Reichsamt (1940c). Weitere Teilergebnisse der Volks- und Berufszählung vom 17.Mai 1939. Die ständige Bevölkerung nach Geschlecht, Alter und Familienstand. Wirtschaft und Statistik, vol. 20, issue 16.

Statistisches Reichsamt (1940d). Weitere Teilergebnisse der Volks- und Berufszählung vom 17.Mai 1939. Die ständige Bevölkerung nach Geschlecht, Alter und Familienstand. Wirtschaft und Statistik, vol. 20, issue 17.

Statistisches Reichsamt (1940e). Weitere Teilergebnisse der Volks- und Berufszählung vom 17.Mai 1939. Die ständige Bevölkerung nach Geschlecht, Alter und Familienstand. Wirtschaft und Statistik, vol. 20, issue 19.

(IV) Figure 5 (Sex ratios in 1946)

Ausschuss der Deutschen Statistiker für die Volks- und Berufszählung 1946 (1949b). Volks- und Berufszählung vom 29. Oktober 1946 in den vier Besatzungszonen und Gross-Berlin. Volkszählung: Tabellenteil. Berlin and Munich: Duncker & Humblot.

(V) Figure 6 (Sex ratios on the state level in 1946 with and without expellees)

Bayerisches Statistisches Landesamt (1948). Statistisches Jahrbuch für Bayern 1947, vol. 23.

Statistisches Landesamt Nordrhein-Westfalen (1950). Landes- und Kreisergebnisse aus der Volks- und Berufszählung 1946, Beiträge zur Statistik des Landes Nordrhein-Westfalen, Issue 3.

Statistisches Zentralamt für die sowjetische Besatzungszone (1949). Volks - und Berufszählung vom 29. Oktober 1946 in der sowjetischen Besatzungszone. Band III: Landes- und Kreisstatistik, issues 1-5.

(d) Reference for map

The shapefile for the map was kindly provided to us by John O'Loughlin. It was created for the following publication: John O'Loughlin. The Electoral Geography of Weimar Germany: Exploratory Spatial Data Analysis (ESDA) of Protestant Support for the Nazi Party. Political Analysis 10, 3 (2002), 217-243.
The shapefile is available from: http://www.colorado.edu/ibs/intdev/johno/pubs.html

3.A.2 Election data

All data on the federal elections of 1983 and 1987 in West Germany and on the federal elections of 1990 and 1994 in reunified Germany was obtained from the Federal Returning Officer: Der Bundeswahlleiter (2018), Der Bundeswahlleiter (2016a), Der Bundeswahlleiter (2016b) and Der Bundeswahlleiter (2017a), respectively; these data are © Der Bundeswahlleiter, Wiesbaden 2016, 2017, 2018.

Data on the members of the 10th East-German *Volkskammer* was obtained from the website Datenbank VOLKPARL (n.d.).

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

with Michael Lechner, Tim Pawlowski and Ute Schüttoff[1]

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 data base 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 codes: Z20, H72, H75, C21

[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.

4.1 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).[2] Since physical inactivity has been shown to increase the risk of suffering from Noncommunicable Diseases (NCD)[3] 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.[4]

Despite the large amount of money spent to supply sports facilities (e.g., German municipalities spent around 4.3 Billion EUR in 2006)[5], little is known whether these public expenditures are effective in fostering sports participation. Only few studies have

[2]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).

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

[4]The precise text is as follows: *"Creation and preservation of built and natural environments which support physical activity in schools, universities, workplaces, clinics and hospitals, and in the wider community, with a particular focus on providing infrastructure to support active transport, i.e. walking and cycling, active recreation and play, and participation in sports."* (World Health Organization, 2013, p. 34).

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

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*.[6] 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.[7] This brief overview suggests that evidence regarding the effect of local sports facility expenditures on sports participation is generally scarce. Moreover, the 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

[6] 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.

[7] 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.

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.[8]

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 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 survey for Germany (Wagner et al., 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.[9] In this paper, we use the approach proposed by Flores et al. (2012), since it allows us to estimate the dose-response function

[8]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.

[9]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.

semi-parametrically in a very flexible and computationally efficient way.[10] 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 researchers who consider using these methodologies in applied work.

Overall, we find no effect of expenditures on sports facilities on the probability to engage in 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 4.3. Results as well as robustness checks are presented in Section 4.4. Sections 4.5 concludes and identifies avenues for future research. Appendices 4.A and 4.B contain more information on the data and descriptive statistics. Appendix 4.C presents further figures and tables with regard to GPS-estimation, common support and balancing tests, Appendix 4.D contains robustness checks and Appendix 4.E additional estimation results.

4.2 Data

4.2.1 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).[11] [12] Our measure of relevant local public expenditures is the sum of net expenditures on

[10] 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.

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

[12] 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 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.

sports facilities (incl. swimming pools) on the municipality level.[13] 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).[14] The expenditure data is combined with regional information from the so-called INKAR database,[15] population and district data (Federal Statistical Office (Destatis), 2002, 2003, 2004, 2005, 2006, 2007) 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.[16]

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, 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.[17] 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

[13] 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.

[14] 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).

[15] 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.

[16] 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.

[17] 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.

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.[18]

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

$$\text{sports-exp}_c = \frac{x}{r}\ \text{sports-exp-pc}_c + \left\{\frac{(r-x)}{r}\left[\frac{\sum w_i^{'}\ \text{sports-exp-neighbor}_i}{\sum w_i^{'}\ \text{Population}_i}\right]\right\} \tag{4.1}$$

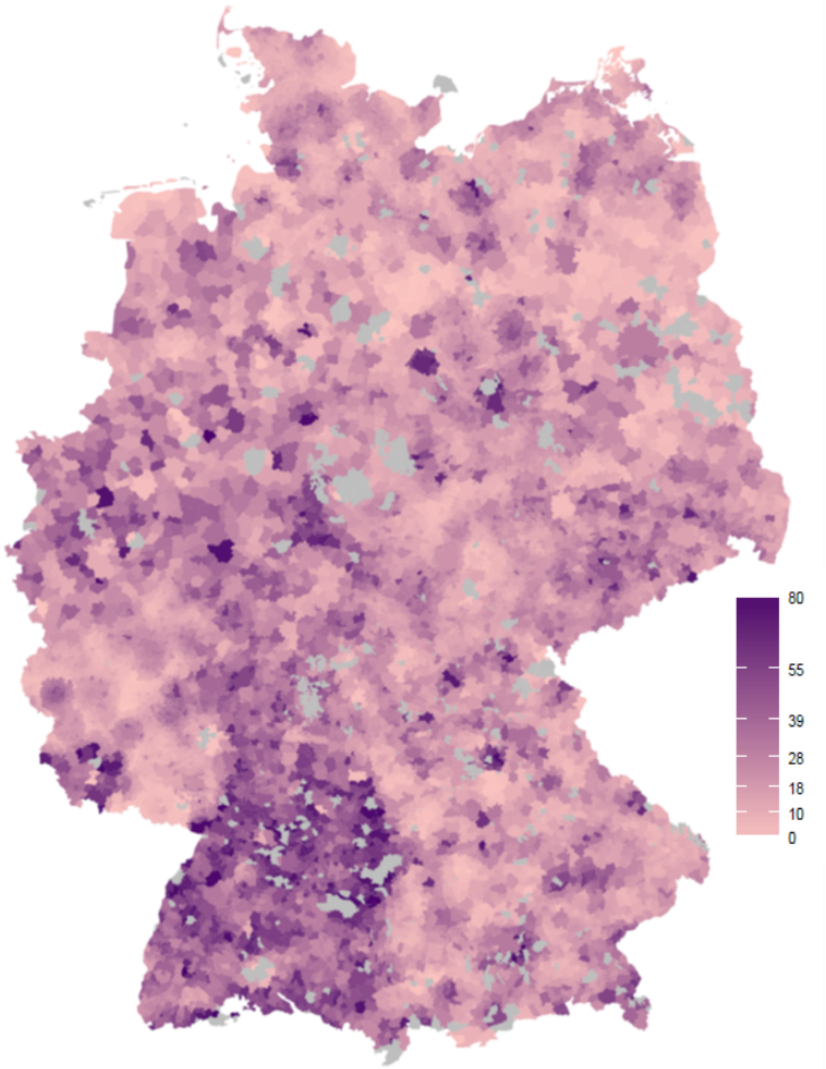
where x is the distance between the centroid of the closest neighboring municipality and the centroid of the municipality itself. r reflects the radius taken into account and is depending on the category of the municipality either 11.3 km, 13.4 km or 14.3 km. Hence, the larger the respective municipality 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 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:

$$w_i^{'} = \frac{w_i}{\sum_{j=1}^{N} w_j} \quad \text{where} \quad w_i = 1/d_i \tag{4.2}$$

Figure 4.1 displays the distance weighted expenditures on sports facilities (averaged over the years 2002-2006) for German municipalities. 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.

[18] 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.

Figure 4.1: 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. The underlying shapefile for this map is provided by ©GeoBasis-DE / BKG 2017. This figure is also contained in Pawlowski et al. (2019).

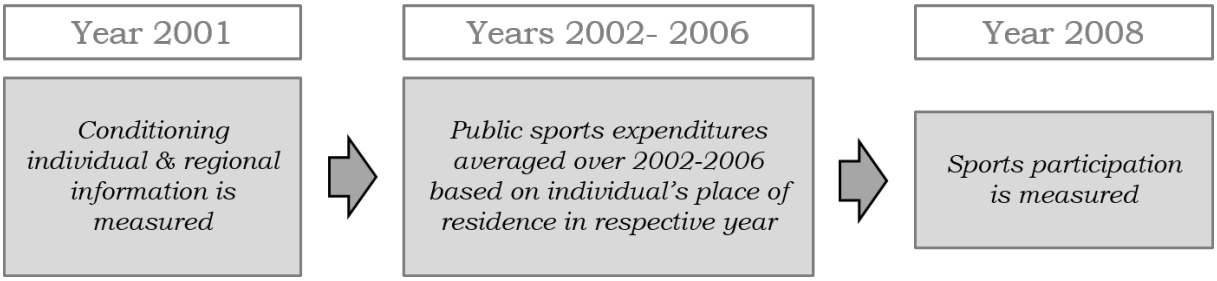
4.2.2 Individual level data: The German Socio-Economic Panel

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).[19] We link the data sets by the official municipality

[19]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.

key *("Amtlicher Gemeindeschlüssel" (AGS))* on an annual basis.[20] Figure 4.2 illustrates how we make use of this panel dataset in the analysis.

Figure 4.2: 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 and regional information from the year 2001 are used as conditioning information, which we will discuss in more detail in Section 4.3.2.

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 engage in sports, we construct a dummy variable *"sports at least monthly"* and study this outcome variable in the main results.[21]

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.[22] As sports expenditures might impact sports

[20] Hence, if an individual lives in Munich from 2002-2004 and from 2005-2006 in Berlin, we assign her expenditure data from Munich for the years 2002-2004 and for the other two years from Berlin.
[21] We also present results for the original variable in a robustness check. The results remain unchanged.
[22] All steps of constructing the SOEP sample are described in Appendix 4.B.

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. Our final sample contains 3,719 individuals.[23] Estimation results for the younger sample (aged 17-30 in 2001) and an older sample (aged 46-65 in 2001) are discussed in Section 4.4.3. 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.[24]

4.2.3 Descriptive statistics

Table 4.1 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.[25] Average per capita facility expenditures in a municipality vary between between 26.7 and 29.2 EUR across years. Figure 4.3 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.

To get an idea how people select into communities with different sports related pub-

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

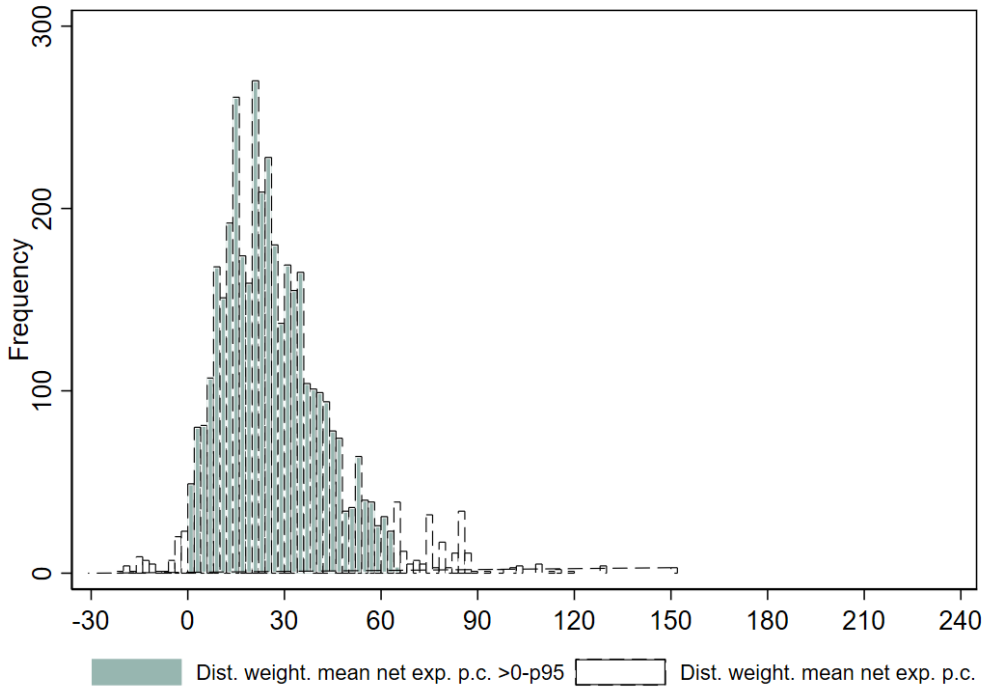
[24] Note that the presented sports participation numbers still include individuals discarded in the last two steps of sample restriction before estimation (described in Appendix 4.B), which encompass restrictions on the sports expenditure variable and individuals pursuing education. In the Eurobarometer 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

Table 4.1: 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 4.3: 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.

participation.

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

lic 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 4.2.[26]

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.

[26]Table 4.2 contains a selection of variables used in the analysis. Descriptive statistics for the full set of variables are presented in Table 4.B.1.

Table 4.2: 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 4.2, own calculations.

4.3 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.

4.3.1 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 $t \epsilon \mathcal{T}$.[27] 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 $t \epsilon \mathcal{T}$ 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):[28]

$$Y_i(t) \perp T_i | X_i \text{ for all t } \epsilon \mathcal{T} \tag{4.3}$$

Hence, conditional on observed pre-treatment variables, the level of treatment T_i is independent 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).

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

[28] 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.

The conditional density of treatment given covariates is denoted by:

$$r(t,x) = f_{T|X}\ (t|X=x) \tag{4.4}$$

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:

$$\begin{gathered}(i)\ \beta(t,r) = E[Y_i(t)|R_i^t = r] = E[Y_i|T_i = t, R_i = r] \\ (ii)\ \mu(t) = E[\beta(t, R_i^t)]\end{gathered} \tag{4.5}$$

4.3.2 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 fol-

lowing 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.[29] 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. Exemplarily, 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.[30] 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

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

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

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 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.

4.3.3 Estimation of the GPS

We apply the two-step estimation procedure proposed by Flores et al. (2012) to estimate the dose-response function semiparametrically.[31] 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.[32] 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" $g(\cdot)$, $g\{E(Y)\} = X\gamma$, 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 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 4.4.1). Parameter estimates are ob-

[31] 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 4.3.4 and to re-estimate the GPS a second time etc.

[32] 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).

tained via maximum likelihood. The log-likelihood function in the case of the Gamma-log model is as follows:[33]

$$\mathcal{L} = \sum_{i=1}^{n} \left\{ \frac{T_i / exp(X_i\gamma) + (X_i\gamma)}{-\phi} + \frac{\phi + 1}{\phi} ln T_i - \frac{ln\phi}{\phi} - ln\Gamma\left(\frac{1}{\phi}\right) \right\} \tag{4.6}$$

The estimated GPS is:[34]

$$\hat{R}_i = \frac{1}{\hat{b}_i^{\hat{a}} \Gamma(\hat{a})} T_i^{\hat{a}-1} e^{-(T_i/\hat{b}_i)} \tag{4.7}$$

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.[35]

4.3.4 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 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:

$$\begin{aligned} CS_q = \Big\{ i : \hat{R}_i^q \, \epsilon \, \Big[max\Big\{ min_{j:Q_j=q} \hat{R}_j^q, min_{j:Q_j \neq q} \hat{R}_j^q \Big\}, \\ min\Big\{ max_{j:Q_j=q} \hat{R}_j^q, max_{j:Q_j \neq q} \hat{R}_j^q \Big\} \Big] \Big\} \end{aligned} \tag{4.8}$$

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.

[33] 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).

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

[35] 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.

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):

$$SD_{before\ Binary}(X) = 100 * \frac{\bar{X}_1 - \bar{X}_0}{\sqrt{\left[V_1(X) + V_0(X)\right]/2}} \tag{4.9}$$

$$SD_{after\ Binary}(X) = 100 * \frac{\bar{X}_{1,M} - \bar{X}_{0,M}}{\sqrt{\left[V_1(X) + V_0(X)\right]/2}} \tag{4.10}$$

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:[36]

$$SD_{before}(X)^{q,q'} = 100 * \frac{\bar{X}^q - \bar{X}^{q'}}{\sqrt{\left[V(X^q) + V(X^{q'})\right]/2}} \tag{4.11}$$

With respect to standardized differences after GPS-adjustment, we start by splitting each of the five GPS predictions into 5 quantiles.[37] 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 to make:

[36]Note that we report mean differences for binary variables.
[37]This first step is hence identical to Hirano and Imbens (2004).

$$SD_{after}(X)^{q,q'} = 100 * \frac{\left[\sum_{c=1}^{C}(\frac{N_c}{N_s}\bar{X}_s^q)\right] - \left[\sum_{c=1}^{C}(\frac{N_c}{N_s}\bar{X}_s^{q'})\right]}{\sqrt{\left[V(X^q) + V(X^{q'})\right]/2}} \tag{4.12}$$

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.[38]

Results of the employed common support procedure are discussed in Section 4.4.1. Table 4.3 summarizes the steps of the proposed common support procedure and indicates the number of observations discarded in each step.

Table 4.3: 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 4.2, own calculations.

Assessment of balancing in this way is appealing for several reasons: First of all, standardized 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.

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

4.3.5 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 $N \to \infty$, $K_h(z) = h^{-1}K(z/h)$.[39] 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:

$$\hat{\mu}(t) = \frac{D_0(t)S_2(t) - D_1(t)S_1(t)}{S_0(t)S_2(t) - S_1^2(t)} \tag{4.13}$$

with $S_j(t) = \sum_{i=1}^{N} \tilde{K}_{h,X}(T_i - t)(T_i - t)^j$ and $D_j(t) = \sum_{i=1}^{N} \tilde{K}_{h,X}(T_i - t)(T_i - t)^j Y_i$, $j = 0,1,2$[40]

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 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).[41]

4.4 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.

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

[40] 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.

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

4.4.1 Estimation of the GPS, common support and balancing tests

Results of the GPS estimation exhibit a similar pattern as the descriptive statistics displayed in Table 4.2: Higher population density, less unemployment and less outward commuting are associated with higher sports facility expenditures (results are displayed in Table 4.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.

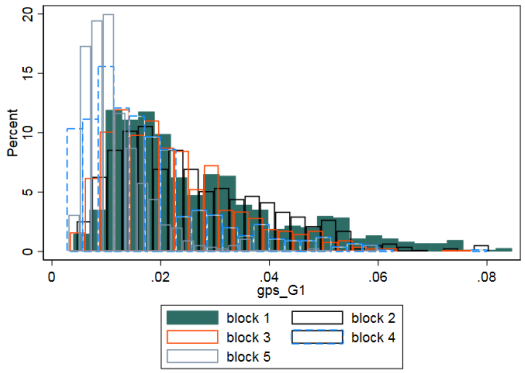
Figure 4.4 presents results from the employed common support procedure. While step 1 and step 3 of the procedure (summarized in Table 4.3) follow Flores et al. (2012), we choose a different graphical representation. Flores et al. (2012) show overlaid histograms of the "own" GPS $\hat{R}_i$ in one block 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 4.4 displays the overlaid histograms of the five blocks at p10 and p30.[42] 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 4.4 contains a selection of pair-wise comparisons of blocks before GPS-adjustment. Large standardized differences can be observed especially for all regional variables.[43]

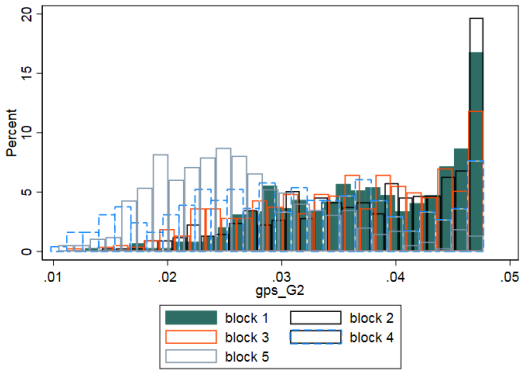
Not surprisingly, these biases are particularly high when we compare blocks "far apart" (for instance Block 1 versus Block 5 or Block 2 versus Block 5). For instance, 18% of individuals in the lowest quintile live in the two southern states, whereas this is the

[42] Figures for the three other points of evaluation can be found in Appendix 4.C.
[43] The full set of all pairwise-comparisons are contained in Table 4.C.2 in Appendix 4.C.

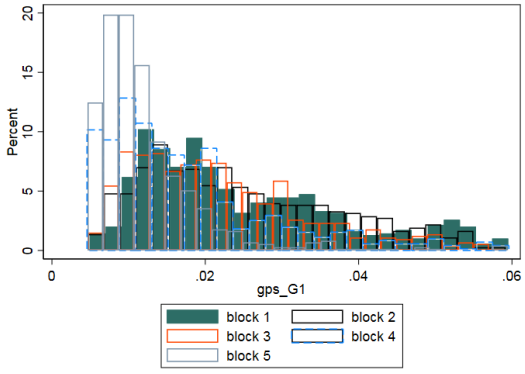
Figure 4.4: Common support



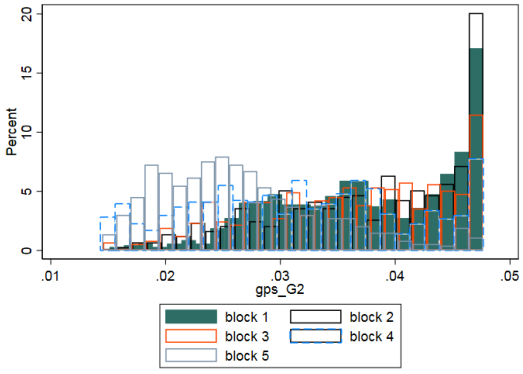
(a) GPS at p10 - before CS condition



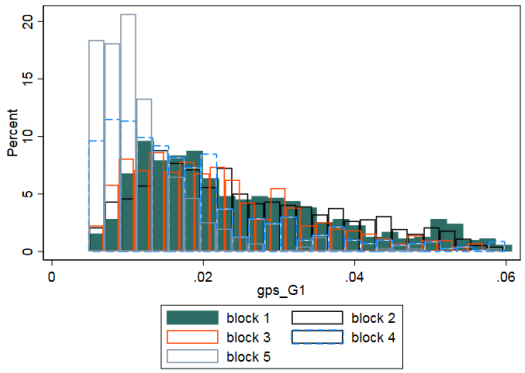
(b) GPS at p30 - before CS condition



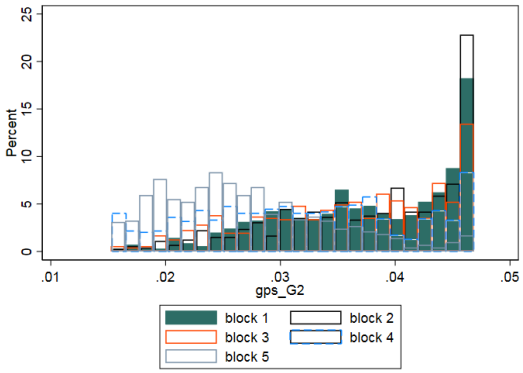
(c) GPS at p10 - after step 1



(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 4.2, own calculations.

Table 4.4: Balancing before GPS adjustment

	Block 1 vs 2			Block 1 vs 5			Block 2 vs 5		
	Mean 1	Mean 2	SD	Mean 1	Mean 5	SD	Mean 2	Mean 5	SD
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 4.2, own calculations.

Table 4.5: 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 4.2, own calculations.

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 4.5. 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.

4.4.2 Effects of sports expenditures on sports participation

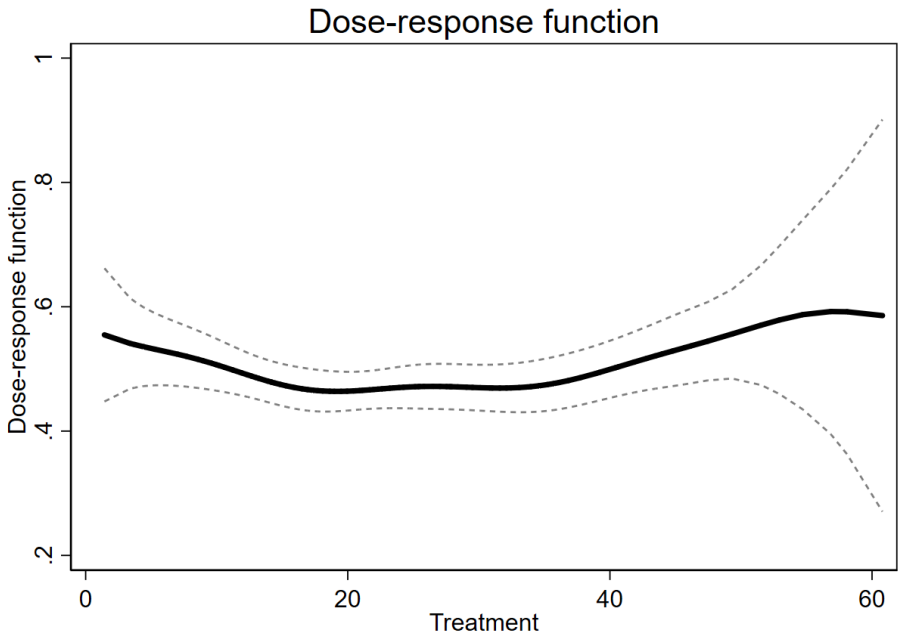
Our main result concerns the dose-response relationship of expenditures on sports facilities on the binary outcome variable *"sports at least monthly"* measured in 2008. A set of robustness exercises is discussed in the next section. Figure 4.5a shows the estimated dose-response function (DRF) and Figure 4.5b the corresponding treatment effect function[44] (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 4.5a. 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:[45] 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 a change in the probability of sports participation by 7.4 percentage points.

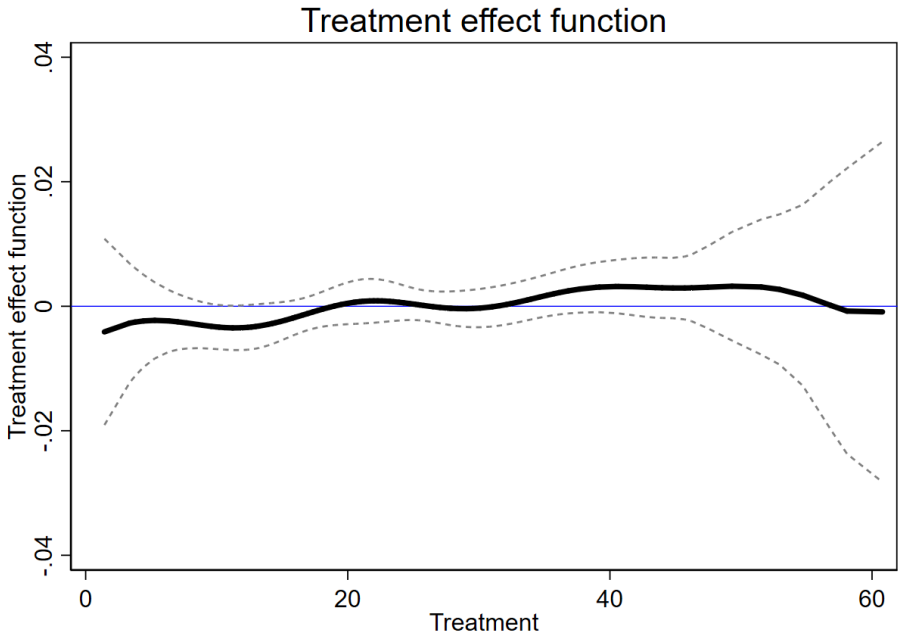
[44] Calculated as $\mu(t + 0.5) - \mu(t)$.

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

Figure 4.5: 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 4.2, own calculations.

4.4.3 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*.[46] 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 4.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 4.D.2 and 4.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 4.D.3) of 0.2 percentage points (compared to 0.6 percentage points in the main specification).[47]

We include lagged sports participation and lagged facility expenditures for reasons discussed in Section 4.3.2 in a next step. Figure 4.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

[46] 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.

[47] 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 4.E.

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 4.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.

4.5 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.

4.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.

4.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[48] (-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

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

Table 4.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 4.2, own calculations.

4.C GPS-estimation, common support & balancing tests

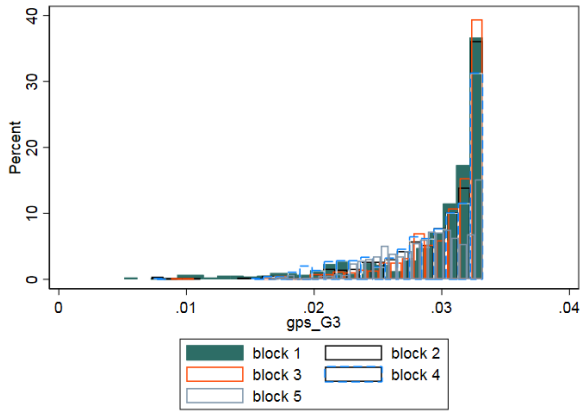
Table 4.C.1: Estimation of GPS

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

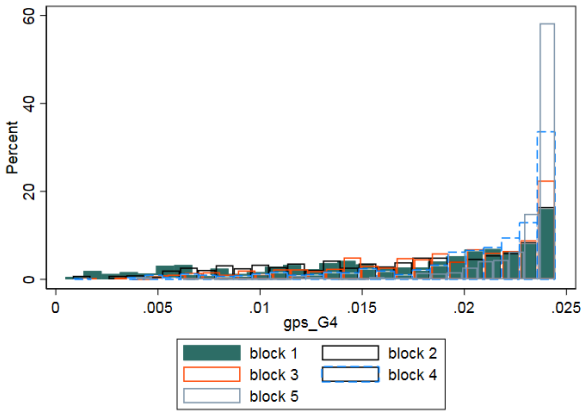
Coefficients of the GPS Model, constant included.

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 4.2, own calculations.

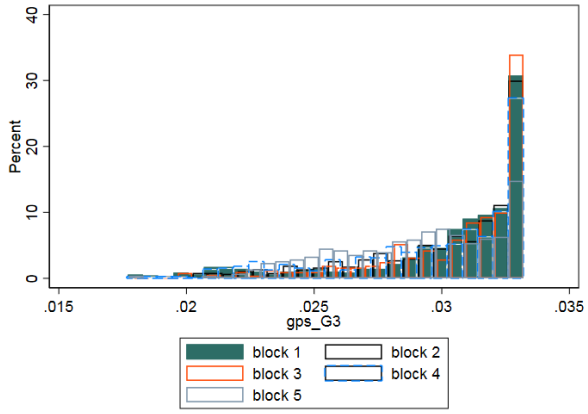
Figure 4.C.1: Common support - A2 sample (main sample)



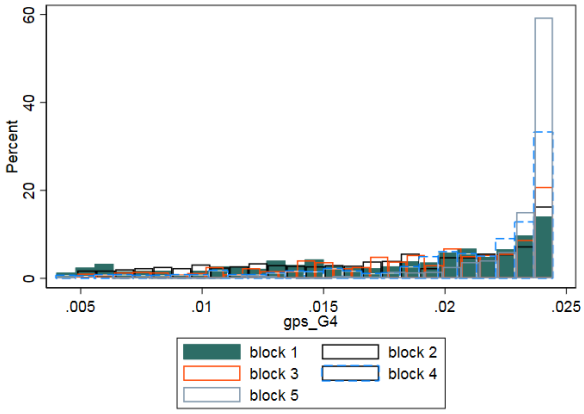
(a) GPS at p50 - before CS condition



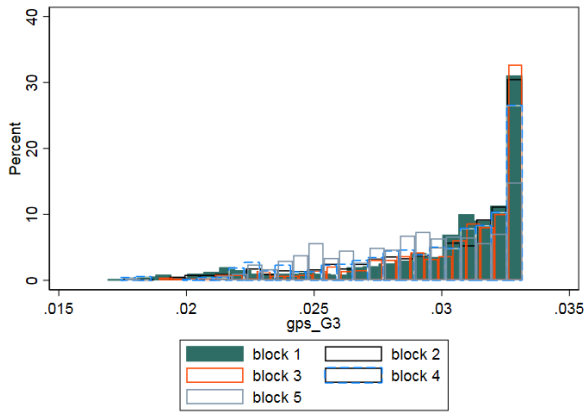
(b) GPS at p70 - before CS condition



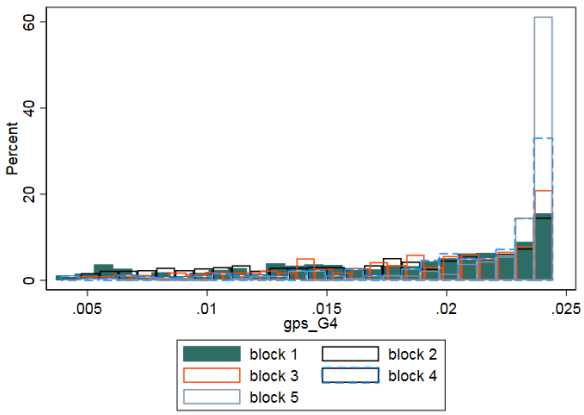
(c) GPS at p50 - after step 1



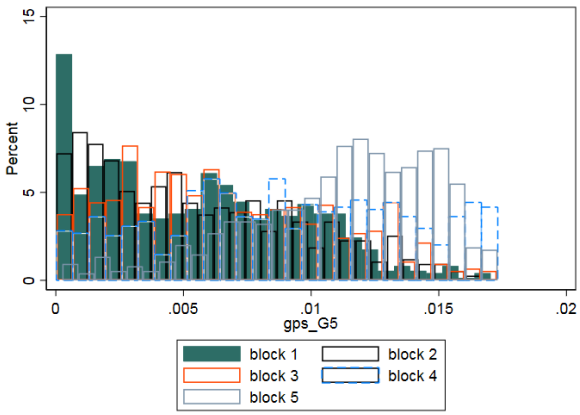
(d) GPS at p70 - after step 1



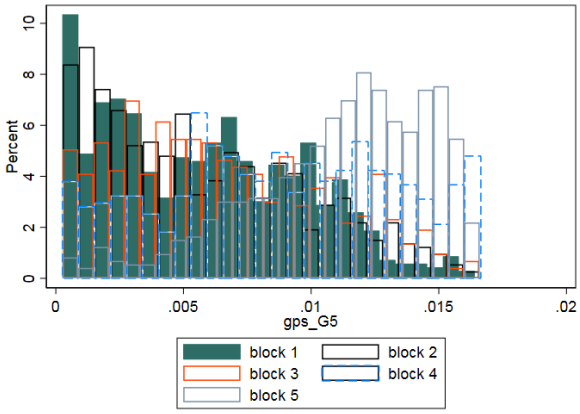
(e) GPS at p50 - after step 3



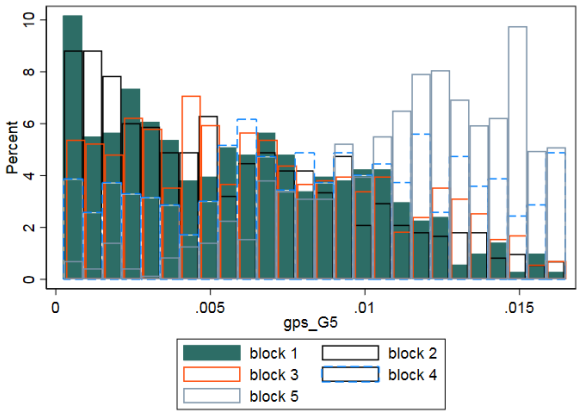
(f) GPS at p70 - after step 3



(g) GPS at p90 - before CS condition



(h) GPS at p90 - after step 1



(i) 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 4.2, own calculations.

Table 4.C.2: Balancing before GPS adjustment

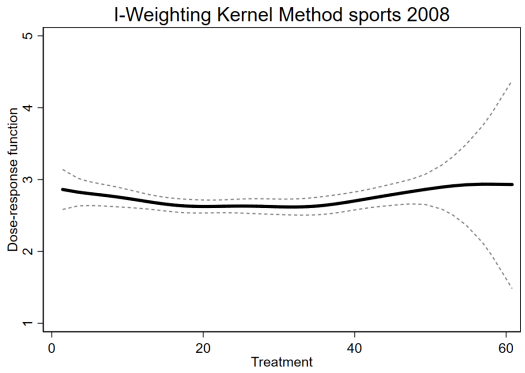
	Block 1 vs 2			Block 1 vs 3			Block 1 vs 4			Block 1 vs 5			Block 2 vs 3		
	Mean 1	Mean 2	SD	Mean 1	Mean 3	SD	Mean 1	Mean 4	SD	Mean 1	Mean 5	SD	Mean 2	Mean 3	SD
ln(population per km2/100)	1.15	1.07	6.54	1.15	1.37	-18.04	1.15	1.71	-43.80	1.15	1.81	-56.21	1.07	1.37	-26.17
ln(average rain/100)	2.02	2.02	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	SD	Mean 2	Mean 5	SD	Mean 3	Mean 4	SD	Mean 3	Mean 5	SD	Mean 4	Mean 5	SD
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

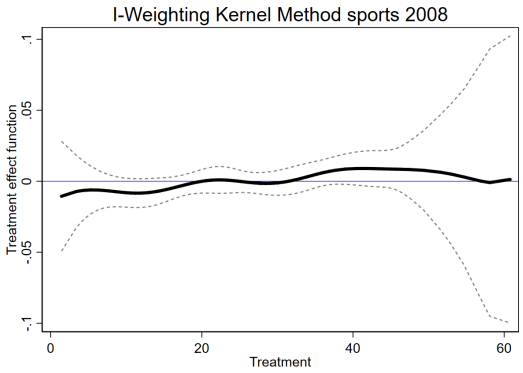
Sources: RDC of the Federal Statistical Office and Statistical Offices of the Länder (2018a; 2018b), SOEP as well as further data sources as discussed in Chapter 4.2, own calculations.

4.D Robustness checks

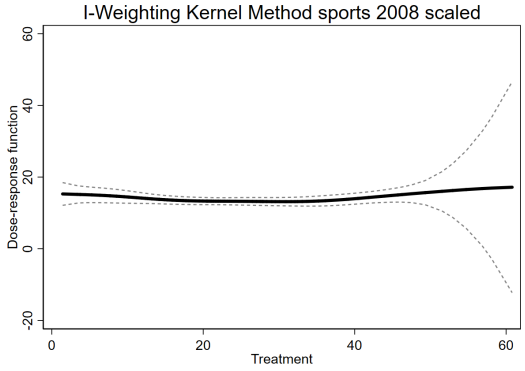
Figure 4.D.1: Results different specifications outcome variable



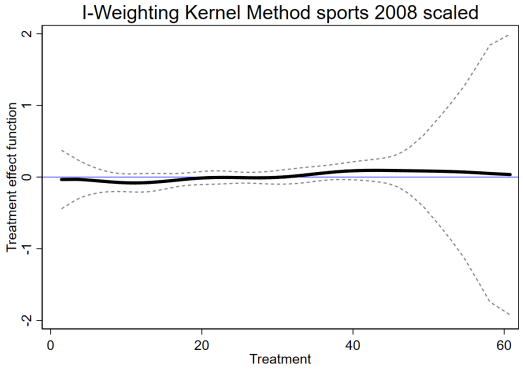
(a) DRF sports 'raw' variable - 2008



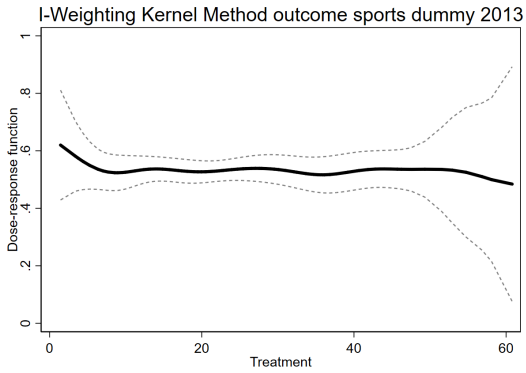
(b) TEF sports 'raw' variable - 2008



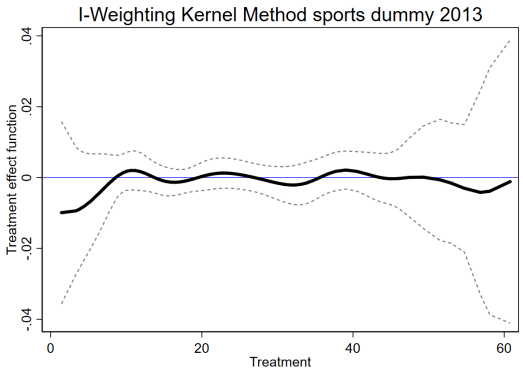
(c) DRF sports variable scaled - 2008



(d) TEF sports variable scaled - 2008



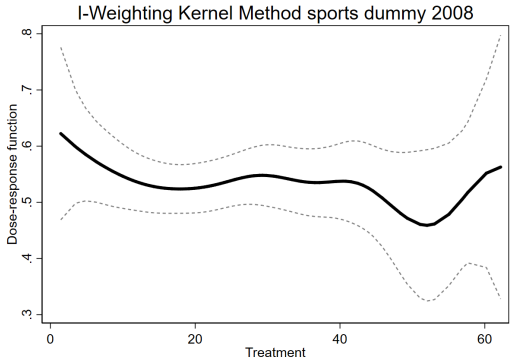
(e) DRF sports dummy - 2013



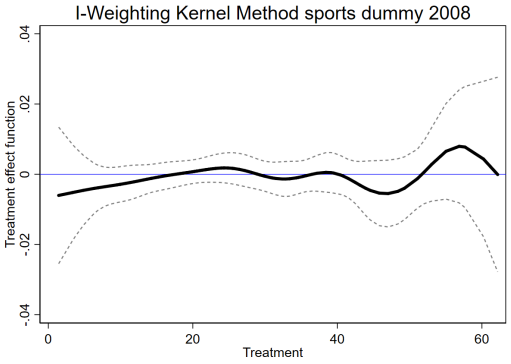
(f) TEF sports dummy - 2013

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 4.2, own calculations.

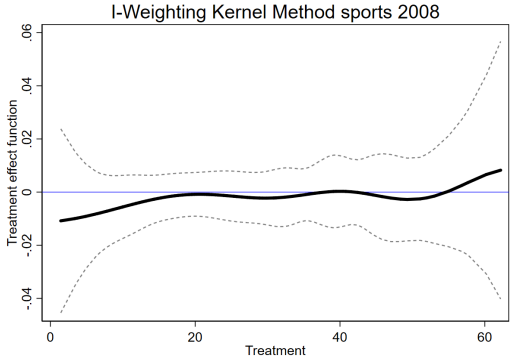
Figure 4.D.2: Results A1 sample (17-30 years of age)



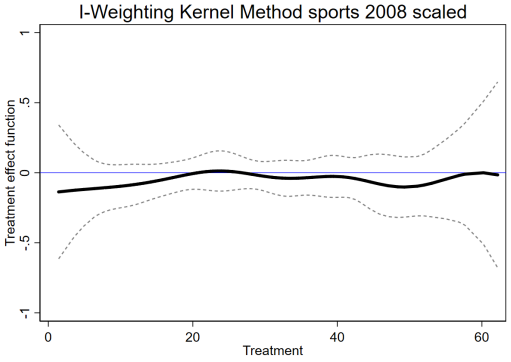
(a) DRF sports dummy - 2008



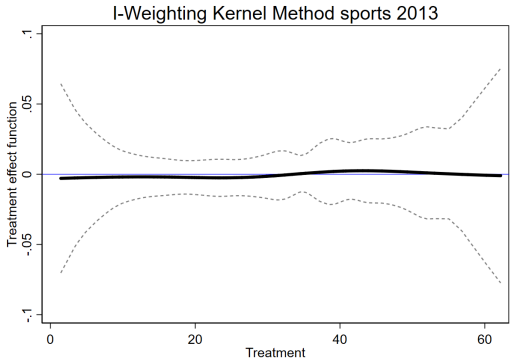
(b) TEF sports dummy - 2008



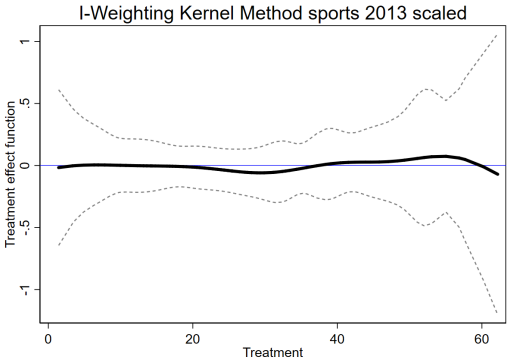
(c) TEF sports 'raw' variable - 2008



(d) TEF sports variable scaled - 2008



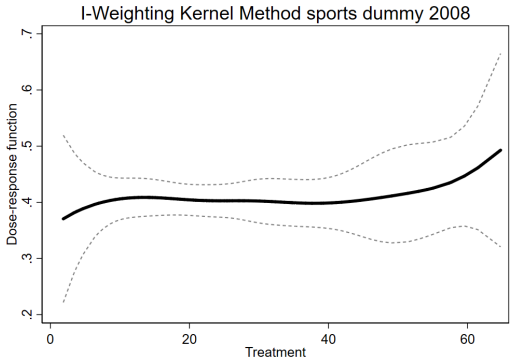
(e) TEF sports 'raw' variable - 2013



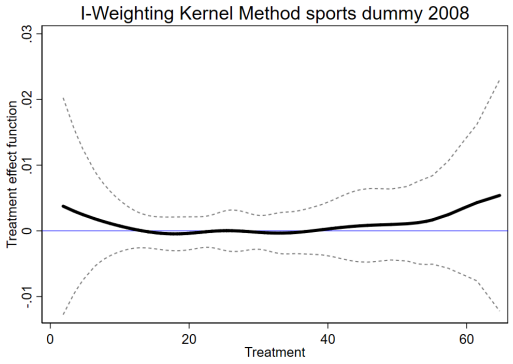
(f) TEF sports variable scaled - 2013

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 4.2, own calculations.

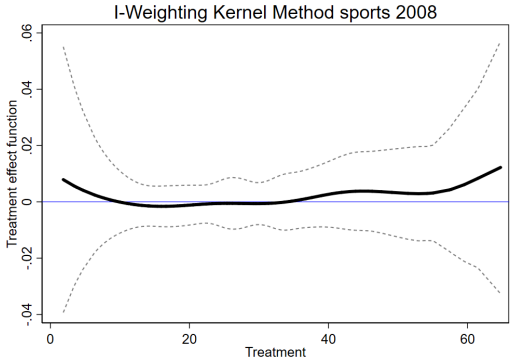
Figure 4.D.3: Results A3 sample (46-65 years of age)



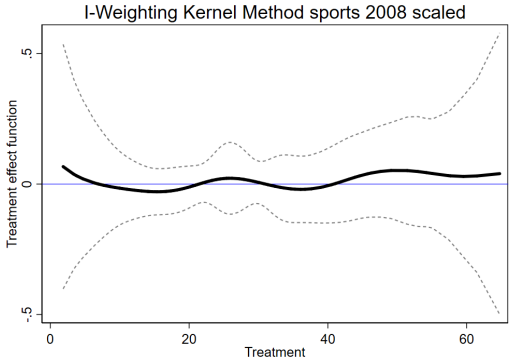
(a) DRF sports dummy - 2008



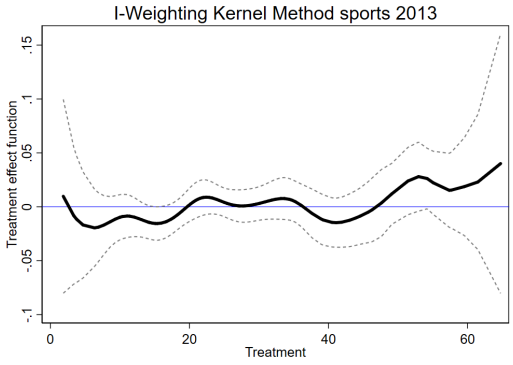
(b) TEF sports dummy - 2008



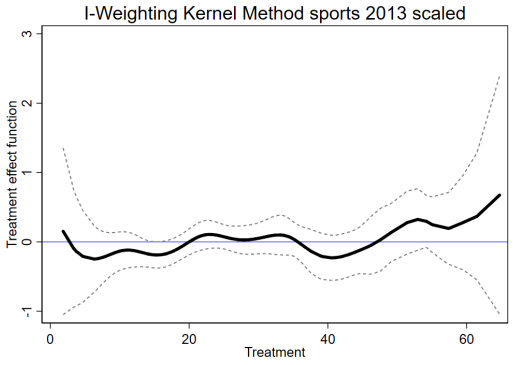
(c) TEF sports 'raw' variable - 2008



(d) TEF sports variable scaled - 2008



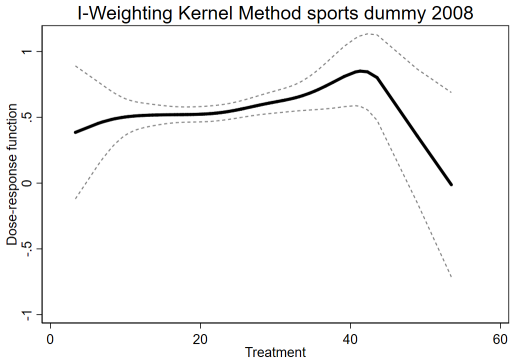
(e) TEF sports 'raw' variable - 2013



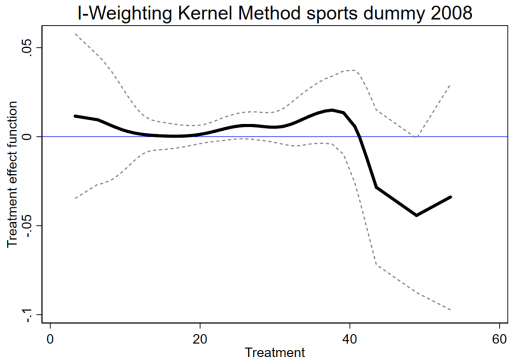
(f) TEF sports variable scaled - 2013

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 4.2, own calculations.

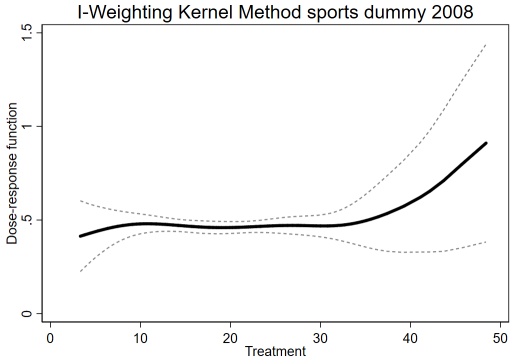
Figure 4.D.4: Results with lagged outcome and treatment



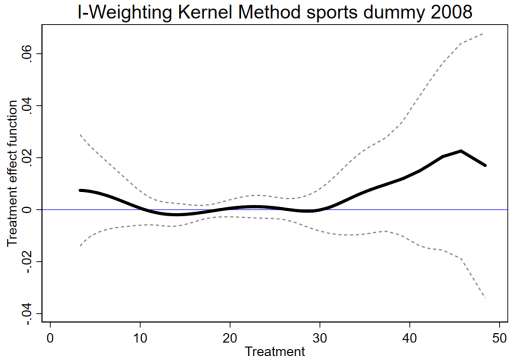
(a) DRF sports dummy 2008 - A1 sample



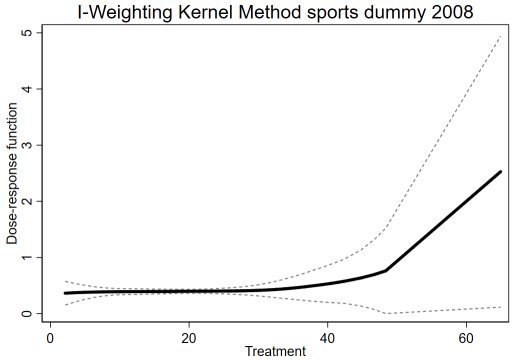
(b) TEF sports dummy 2008 - A1 sample



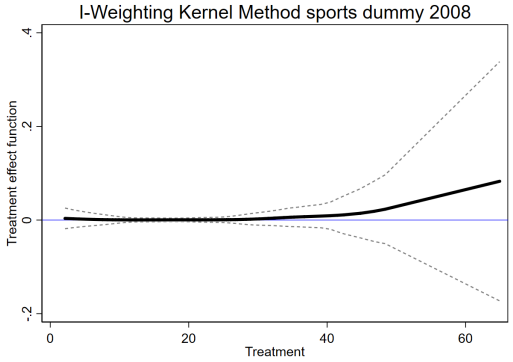
(c) DRF sports dummy 2008 - A2 sample



(d) TEF sports dummy 2008 - A2 sample



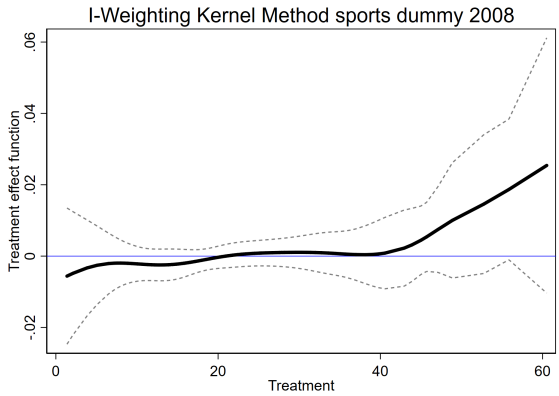
(e) DRF sports dummy 2008 - A3 sample



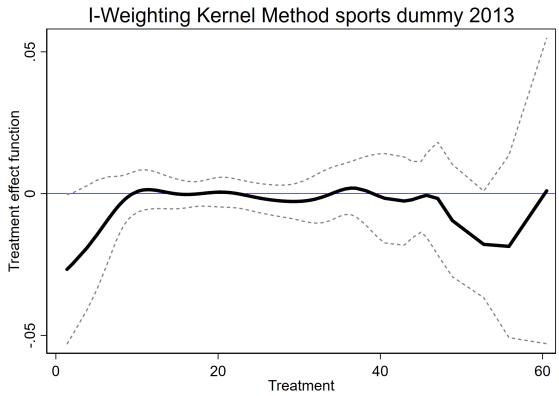
(f) TEF sports dummy 2008 - A3 sample

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 4.2, own calculations.

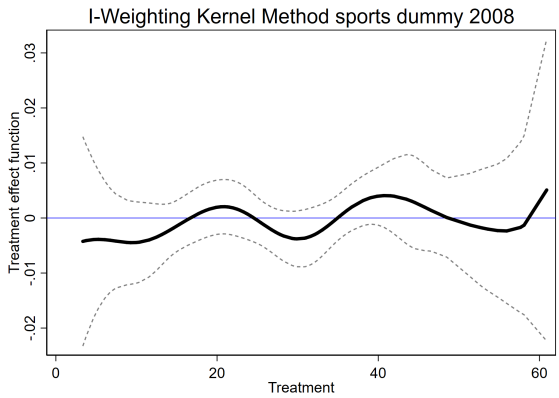
Figure 4.D.5: Results small / large municipalities



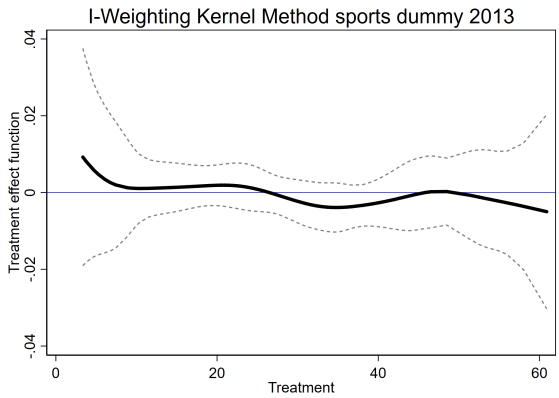
(a) Sports dummy 2008 - Small municipalities



(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 4.2, own calculations.

4.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 4.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 4.2, own calculations.

Table 4.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 4.2, own calculations.

Table 4.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 4.2, own calculations.

Table 4.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 4.2, own calculations.

Table 4.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 4.2, own calculations.

Table 4.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 4.2, own calculations.

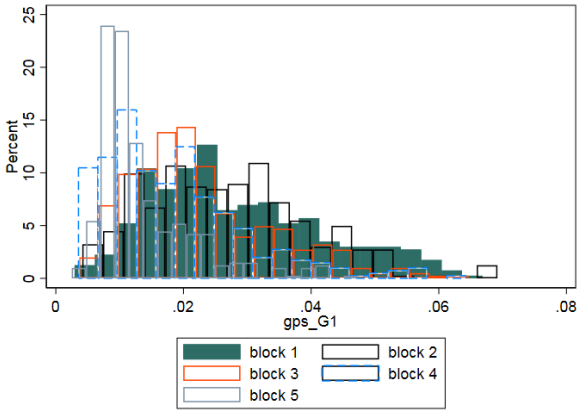
Table 4.E.7: Estimation of GPS - A2 sample - large municipalities

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

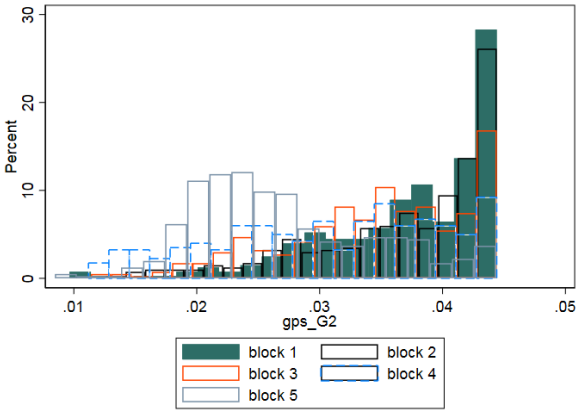
Coefficients of the GPS Model

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 4.2, own calculations.

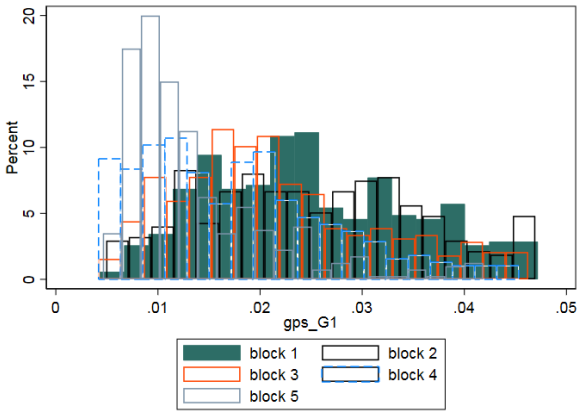
Figure 4.E.1: Common support - A1 sample (17-30 years)



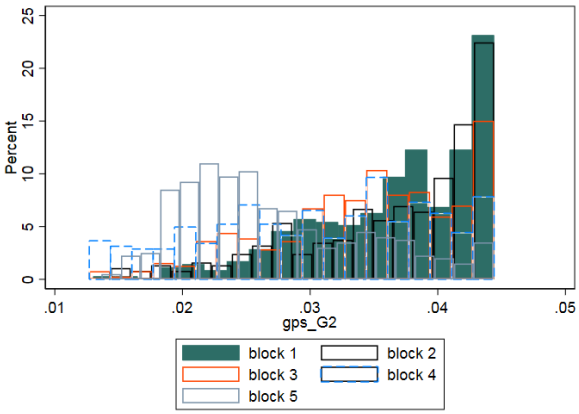
(a) GPS at p10 - before CS condition



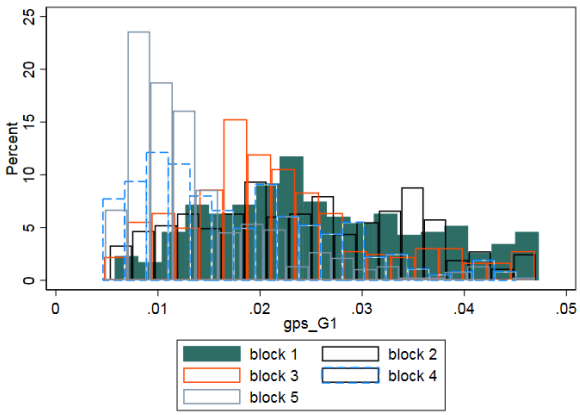
(b) GPS at p30 - before CS condition



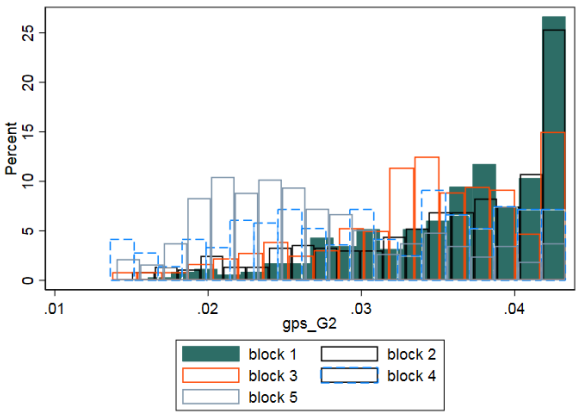
(c) GPS at p10 - after step 1



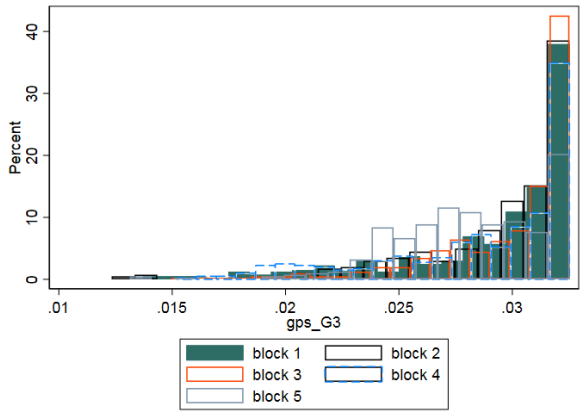
(d) GPS at p30 - after step 1



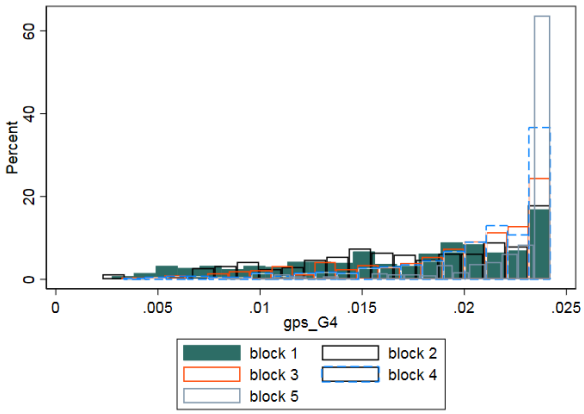
(e) GPS at p10 - after step 3



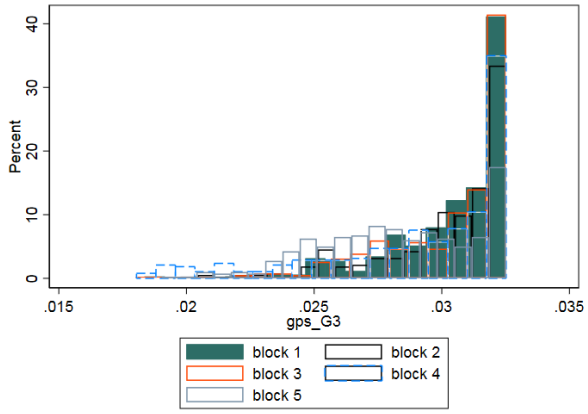
(f) GPS at p30 - after step 3



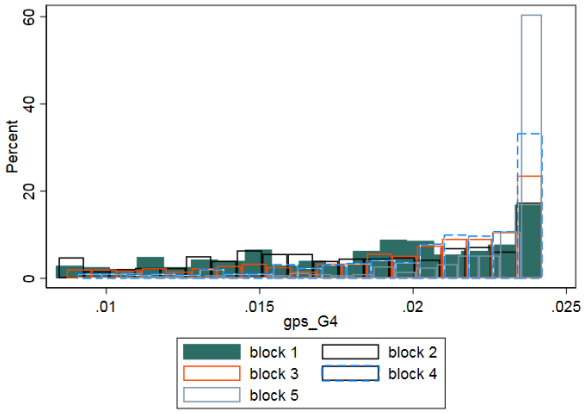
(g) GPS at p50 - before CS condition



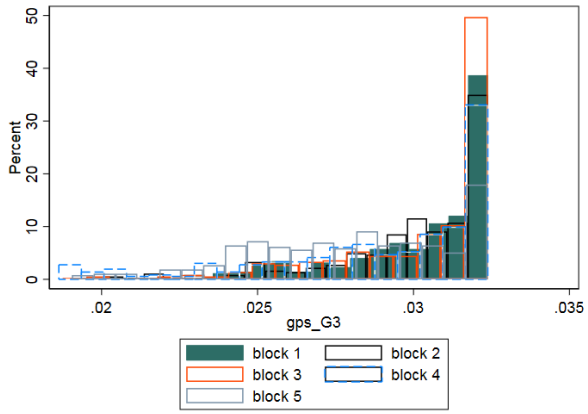
(h) GPS at p70 - before CS condition



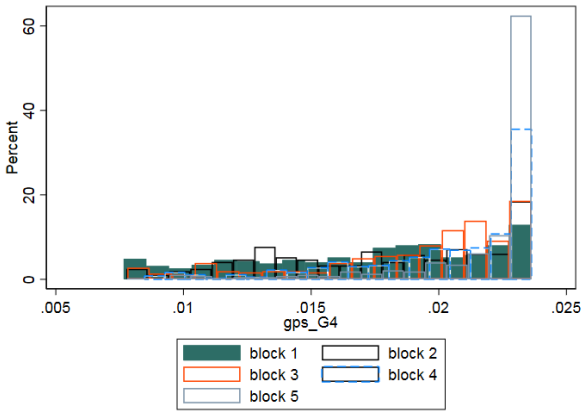
(i) GPS at p50 - after step 1



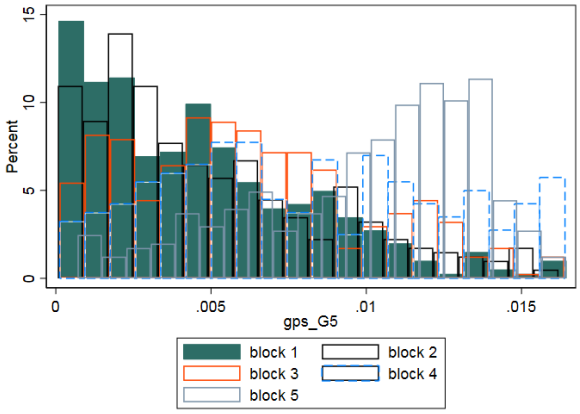
(j) GPS at p70 - after step 1



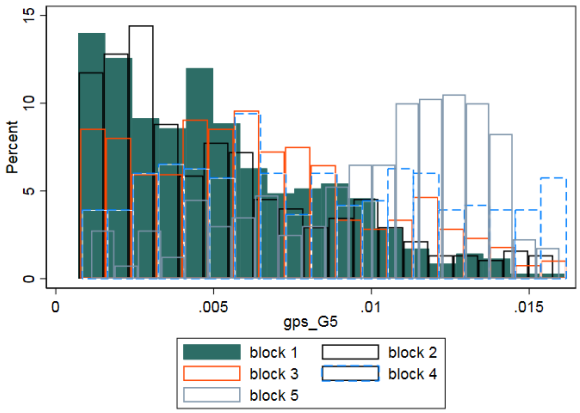
(k) GPS at p50 - after step 3



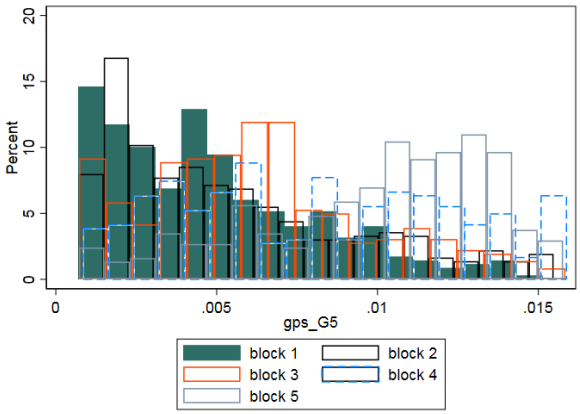
(l) GPS at p70 - after step 3



(m) GPS at p90 - before CS condition



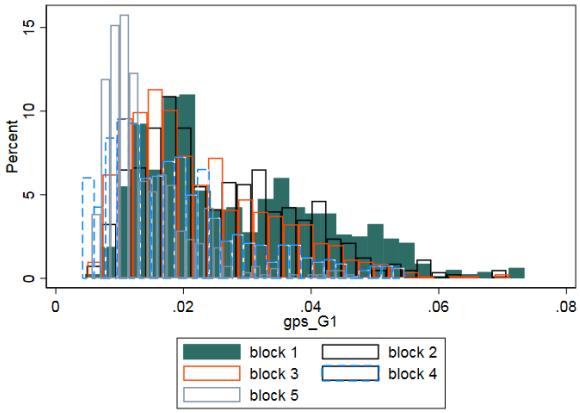
(n) GPS at p90 - after step 1



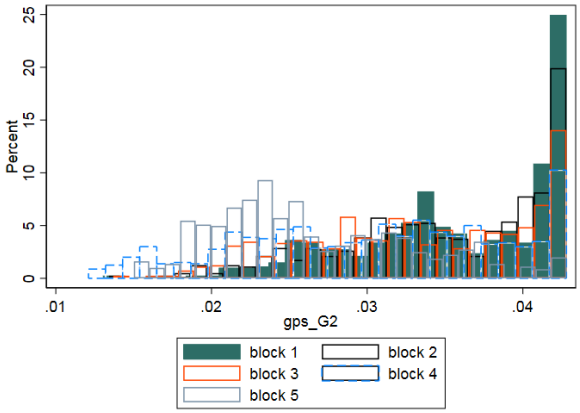
(o) GPS at p90 - after step 3

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 4.2, own calculations.

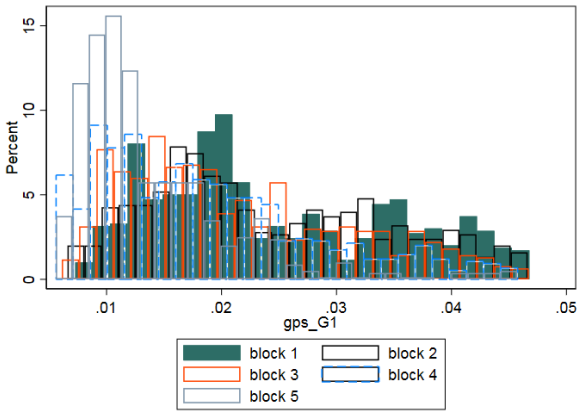
Figure 4.E.2: Common support - A3 sample (46-65 years)



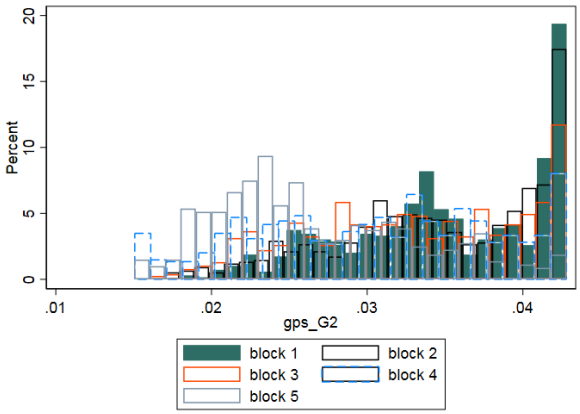
(a) GPS at p10 - before CS condition



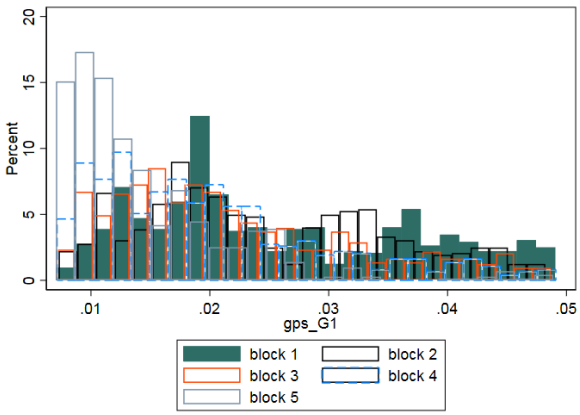
(b) GPS at p30 - before CS condition



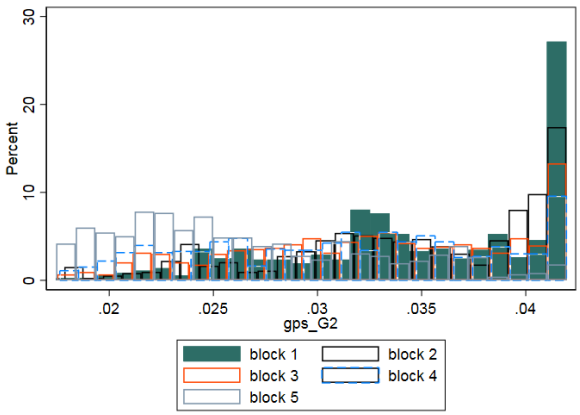
(c) GPS at p10 - after step 1



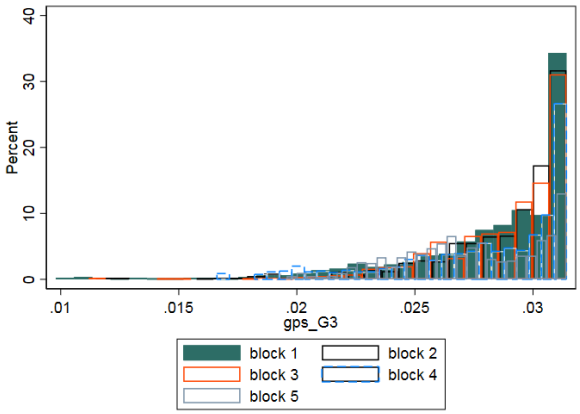
(d) GPS at p30 - after step 1



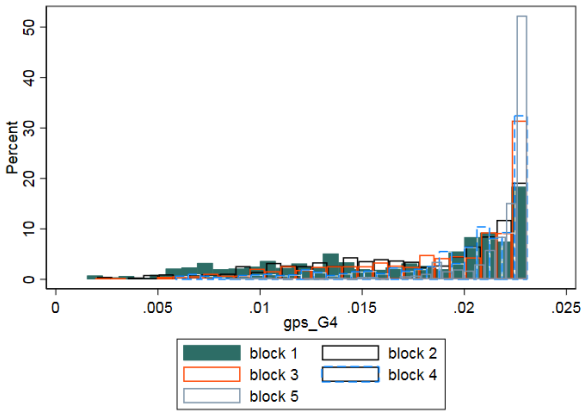
(e) GPS at p10 - after step 3



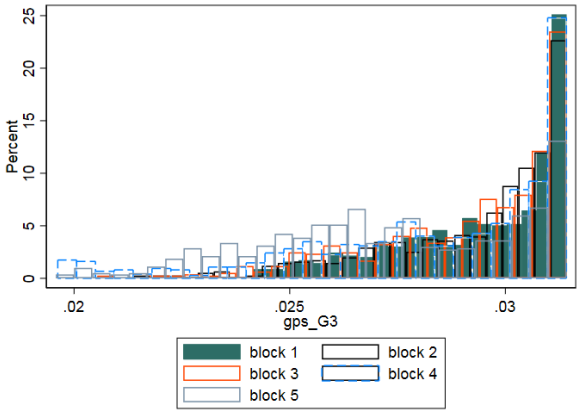
(f) GPS at p30 - after step 3



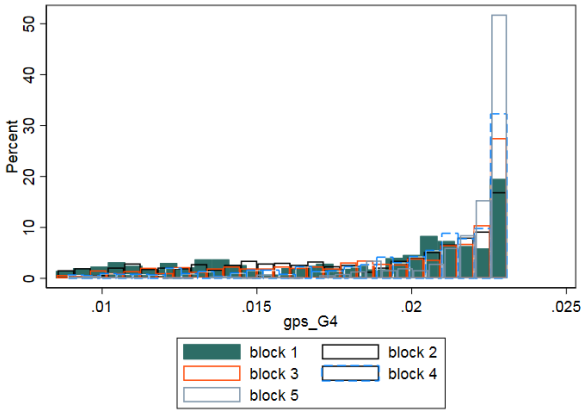
(g) GPS at p50 - before CS condition



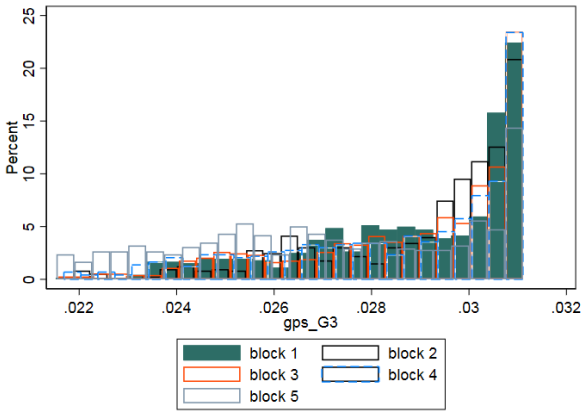
(h) GPS at p70 - before CS condition



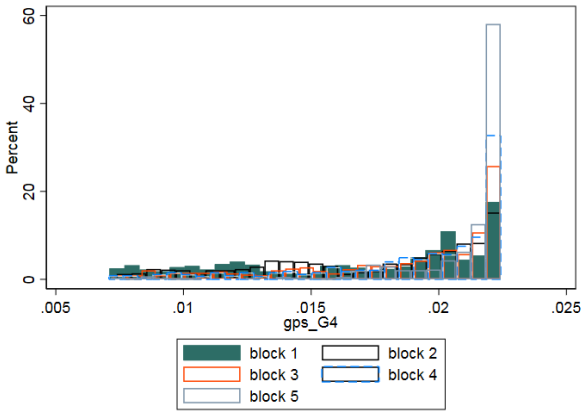
(i) GPS at p50 - after step 1



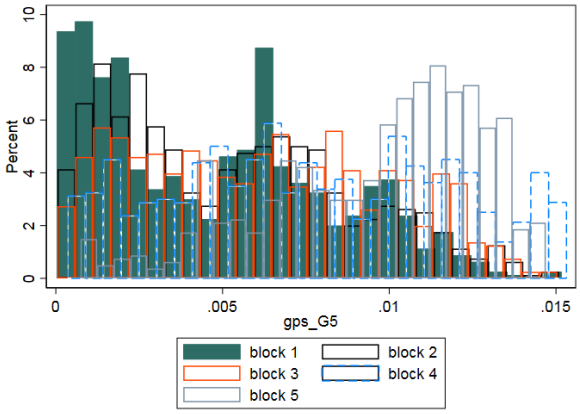
(j) GPS at p70 - after step 1



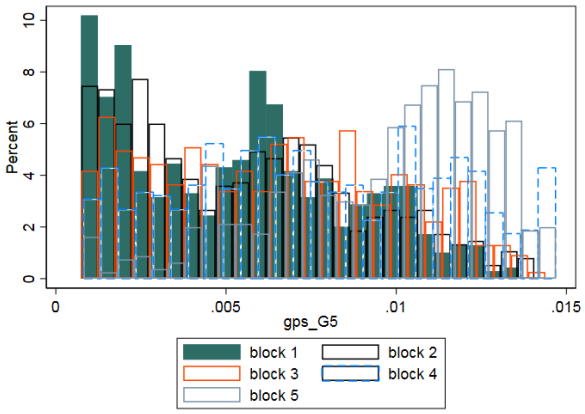
(k) GPS at p50 - after step 3



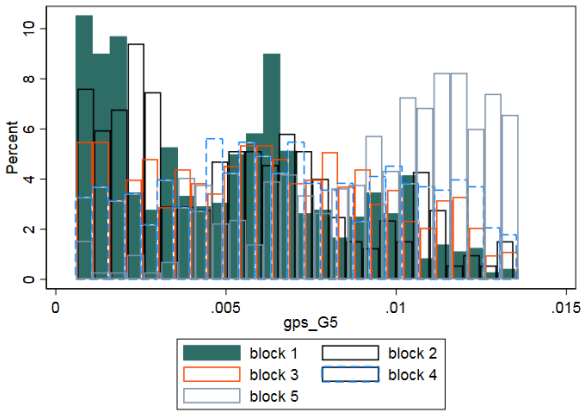
(l) GPS at p70 - after step 3



(m) GPS at p90 - before CS condition



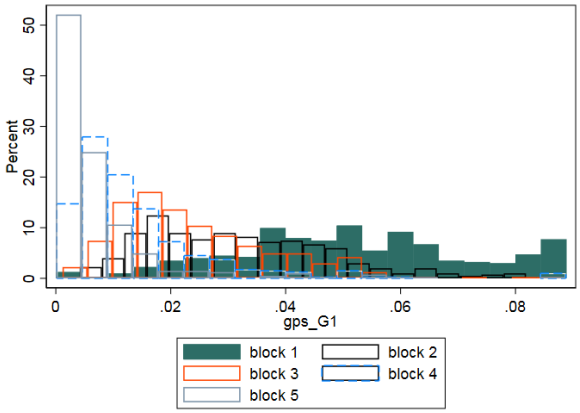
(n) GPS at p90 - after step 1



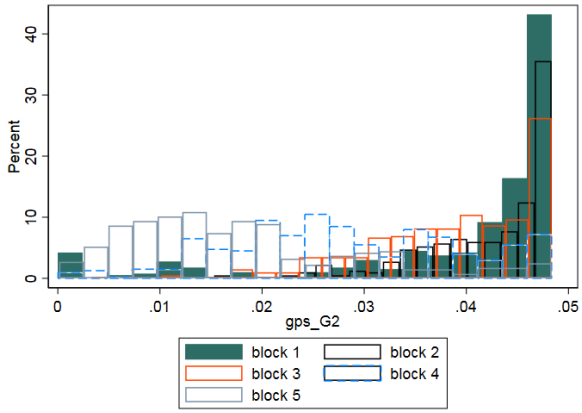
(o) GPS at p90 - after step 3

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 4.2, own calculations.

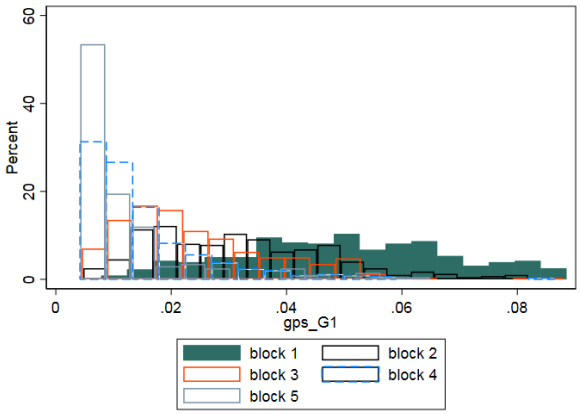
Figure 4.E.3: Common support - A1 with lagged outcome and treatment (17-30 years)



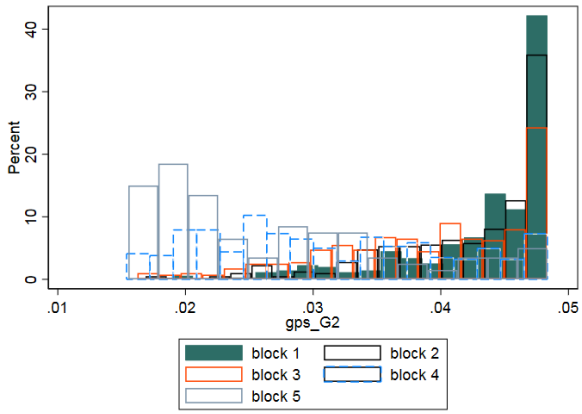
(a) GPS at p10 - before CS condition



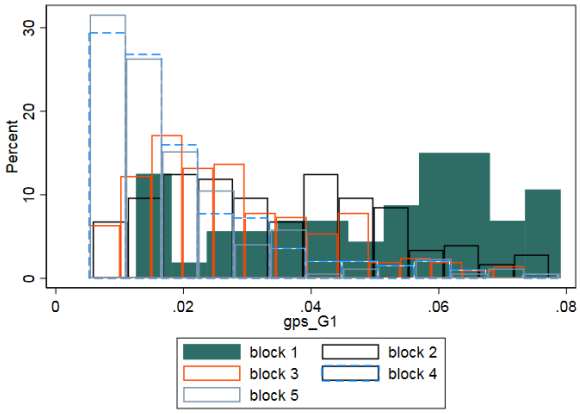
(b) GPS at p30 - before CS condition



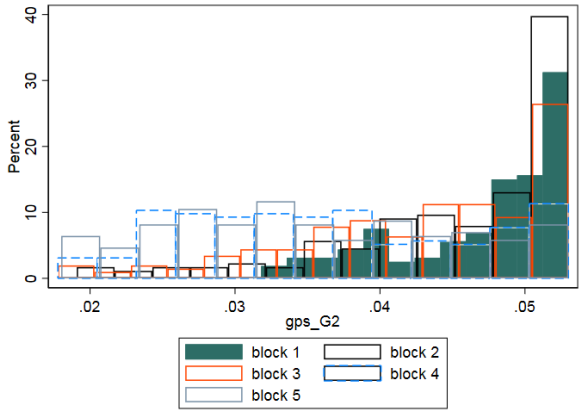
(c) GPS at p10 - after step 1



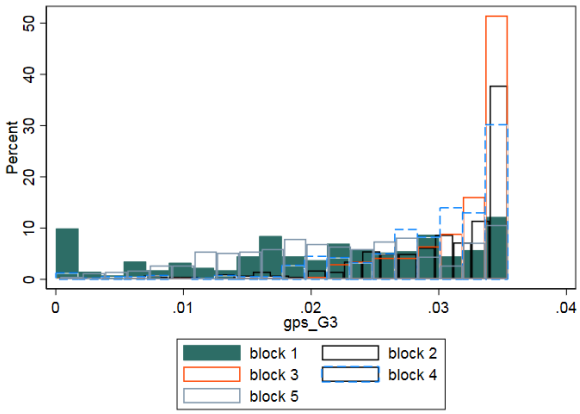
(d) GPS at p30 - after step 1



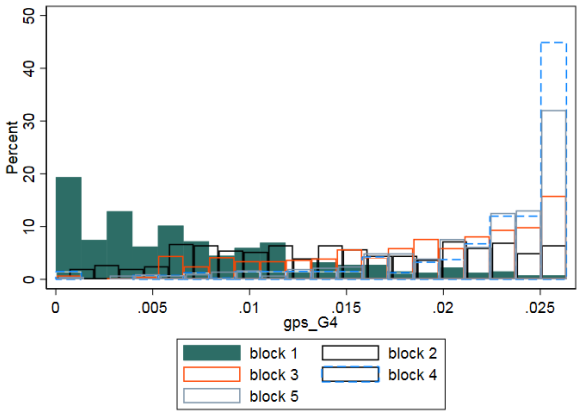
(e) GPS at p10 - after step 3



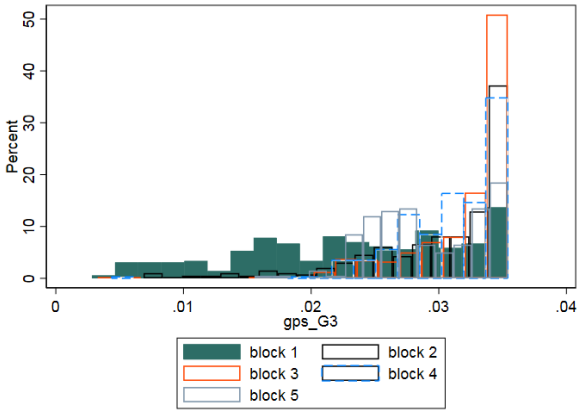
(f) GPS at p30 - after step 3



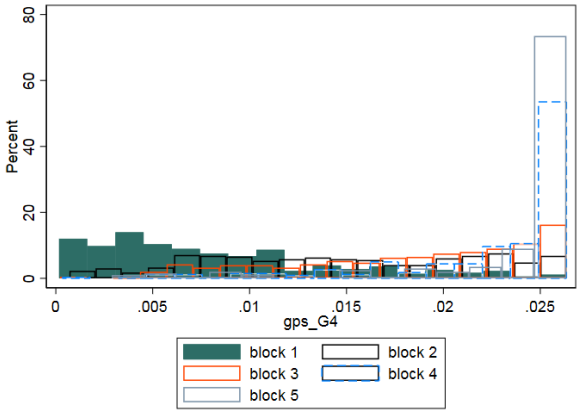
(g) GPS at p50 - before CS condition



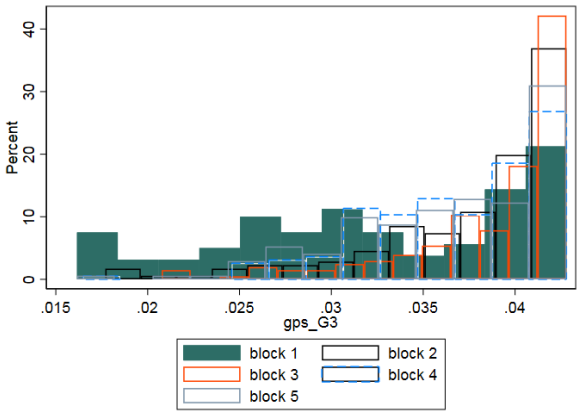
(h) GPS at p70 - before CS condition



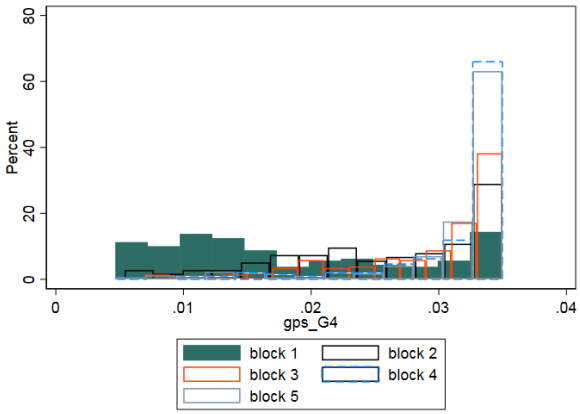
(i) GPS at p50 - after step 1



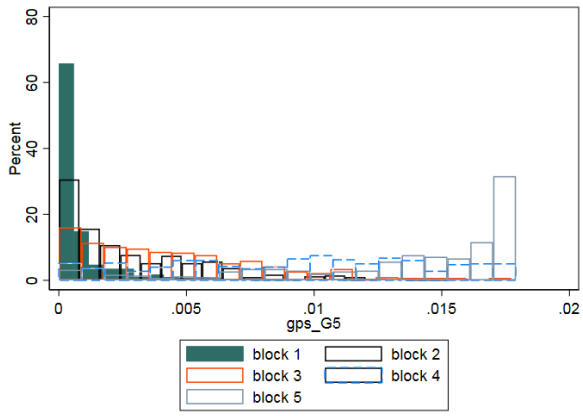
(j) GPS at p70 - after step 1



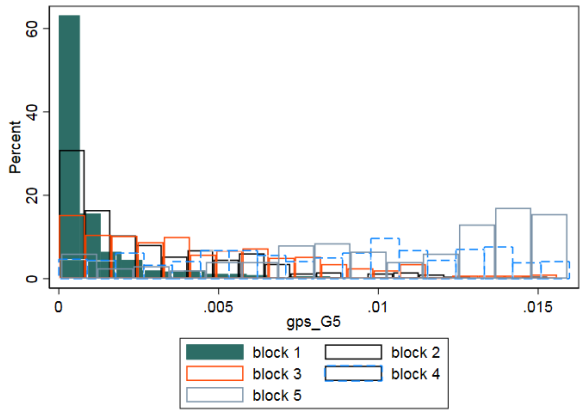
(k) GPS at p50 - after step 3



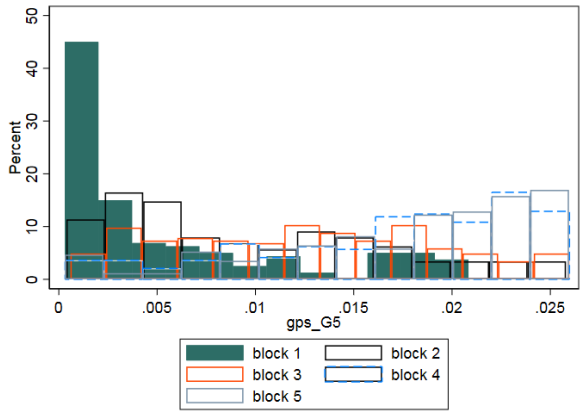
(l) GPS at p70 - after step 3



(m) GPS at p90 - before CS condition



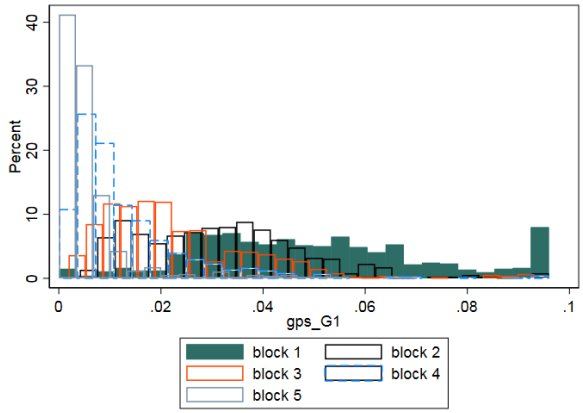
(n) GPS at p90 - after step 1



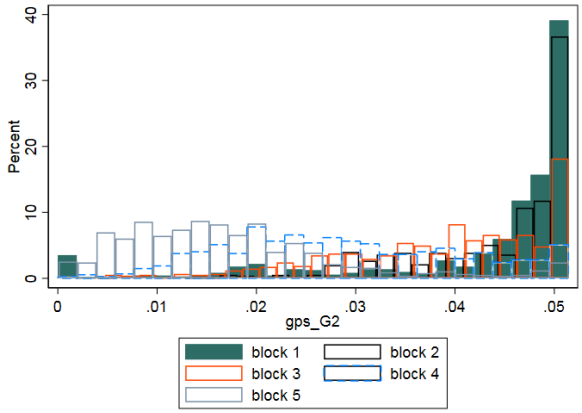
(o) GPS at p90 - after step 3

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 4.2, own calculations.

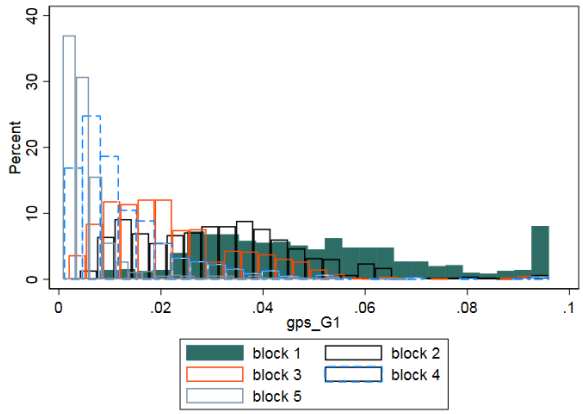
Figure 4.E.4: Common support - A2 with lagged outcome and treatment (31-45 years)



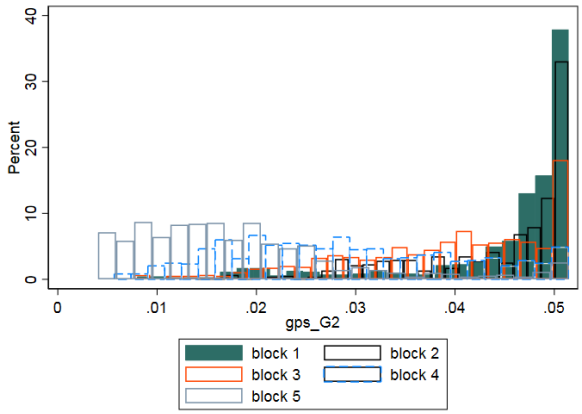
(a) GPS at p10 - before CS condition



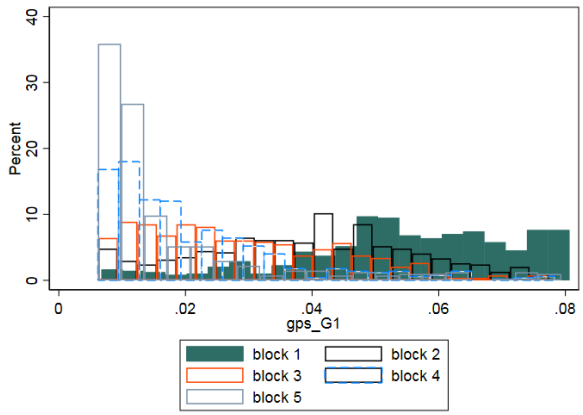
(b) GPS at p30 - before CS condition



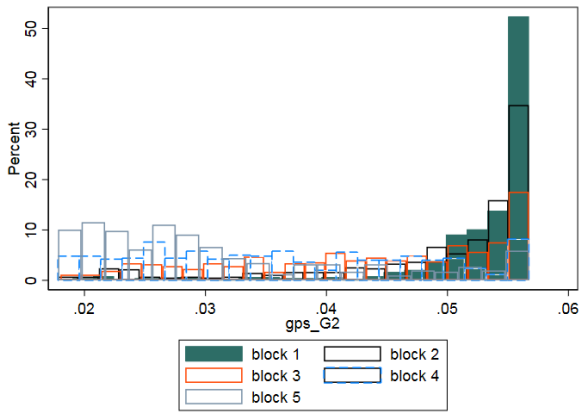
(c) GPS at p10 - after step 1



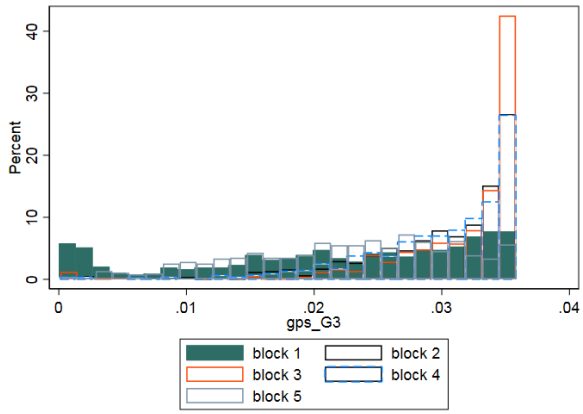
(d) GPS at p30 - after step 1



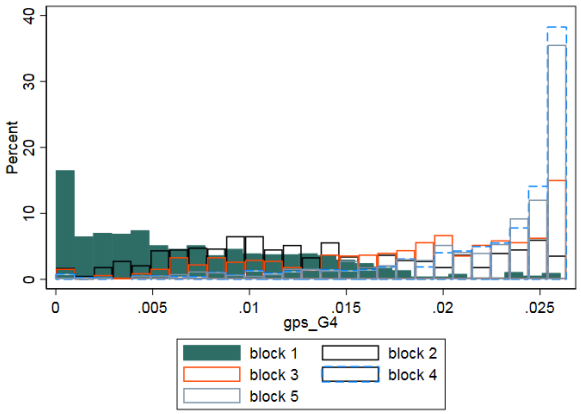
(e) GPS at p10 - after step 3



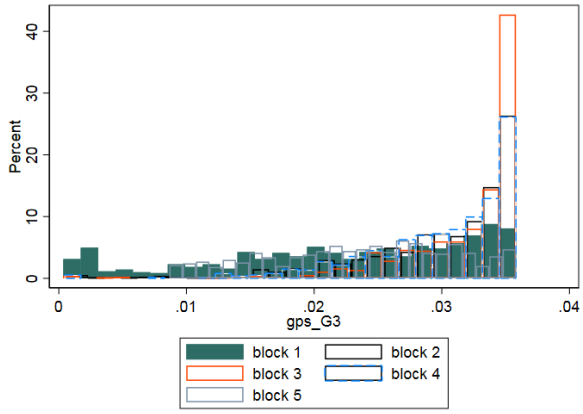
(f) GPS at p30 - after step 3



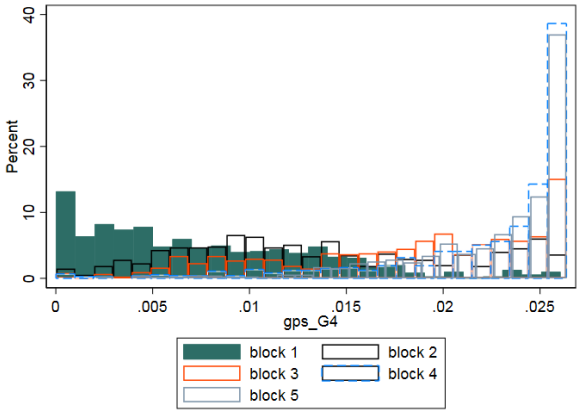
(g) GPS at p50 - before CS condition



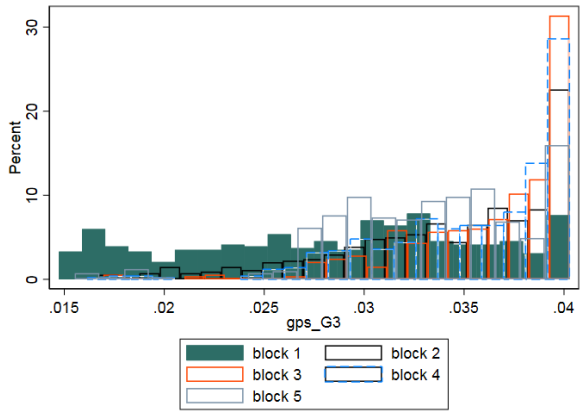
(h) GPS at p70 - before CS condition



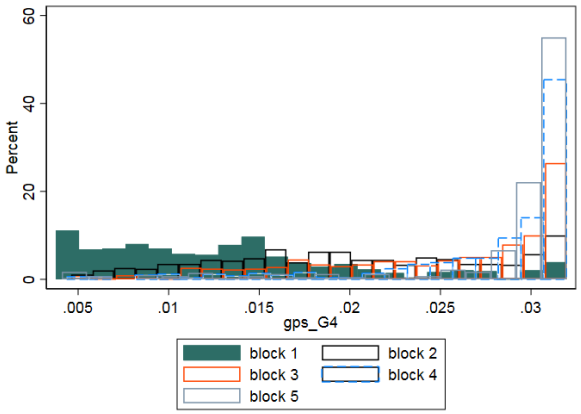
(i) GPS at p50 - after step 1



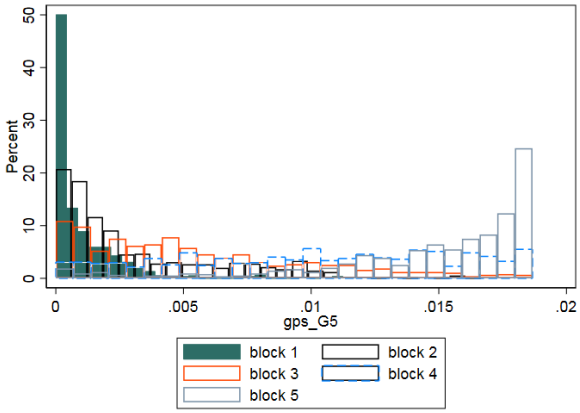
(j) GPS at p70 - after step 1



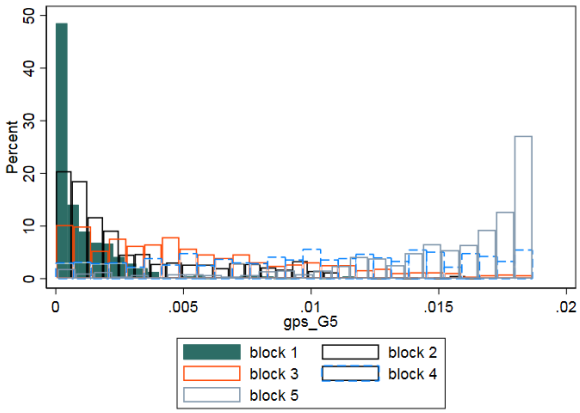
(k) GPS at p50 - after step 3



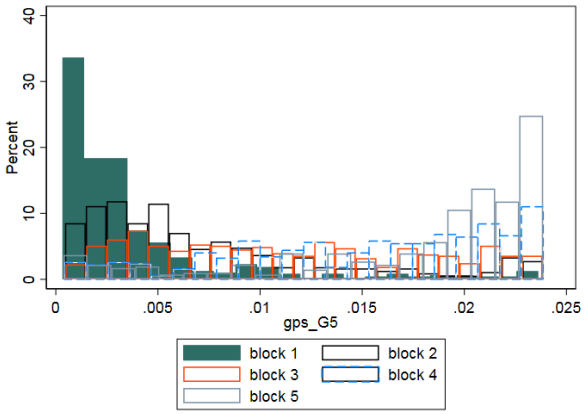
(l) GPS at p70 - after step 3



(m) GPS at p90 - before CS condition



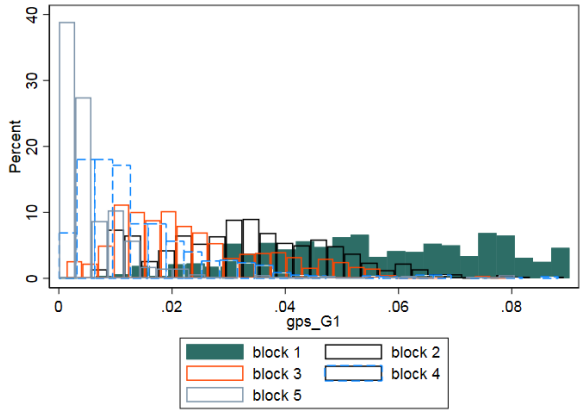
(n) GPS at p90 - after step 1



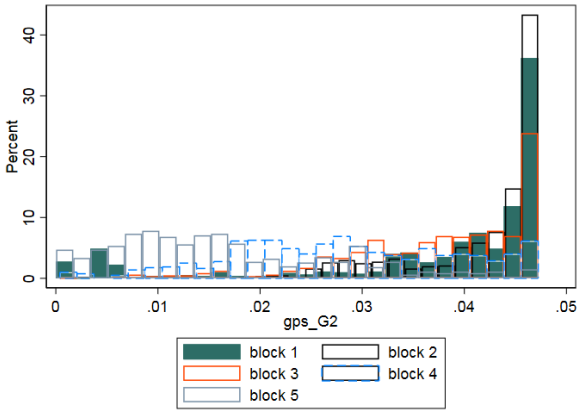
(o) GPS at p90 - after step 3

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 4.2, own calculations.

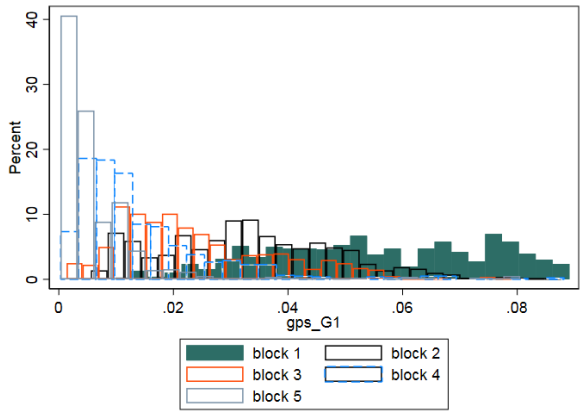
Figure 4.E.5: Common support - A3 with lagged outcome and treatment (46-65 years)



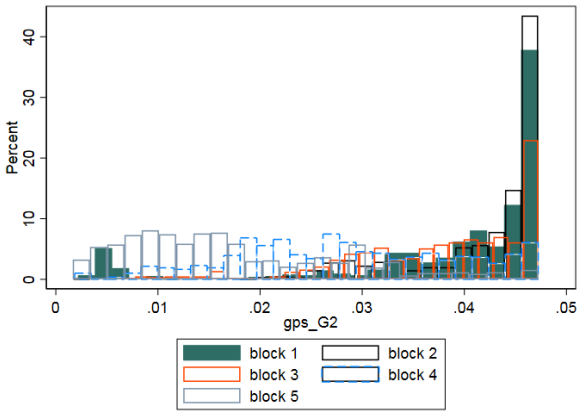
(a) GPS at p10 - before CS condition



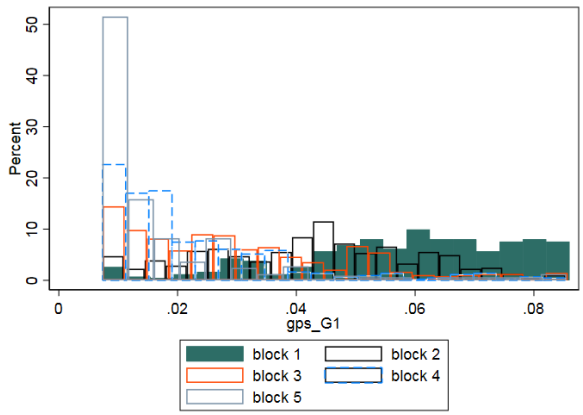
(b) GPS at p30 - before CS condition



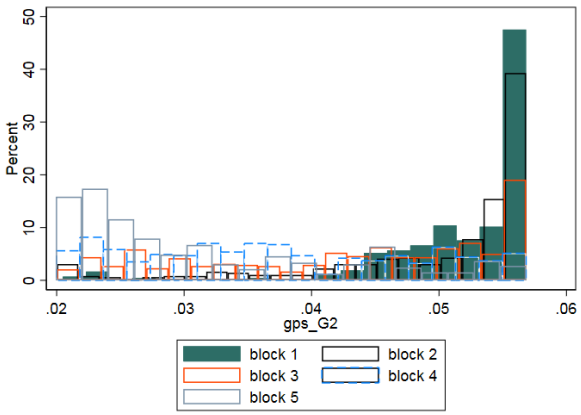
(c) GPS at p10 - after step 1



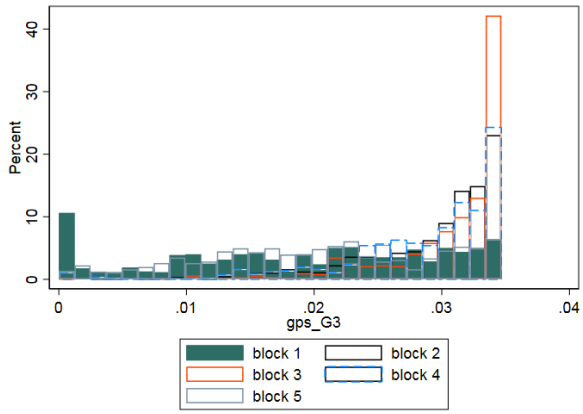
(d) GPS at p30 - after step 1



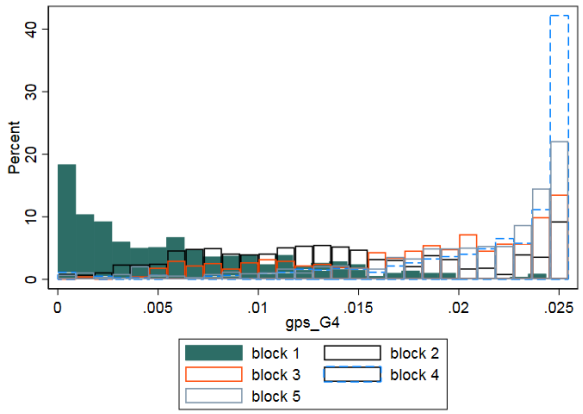
(e) GPS at p10 - after step 3



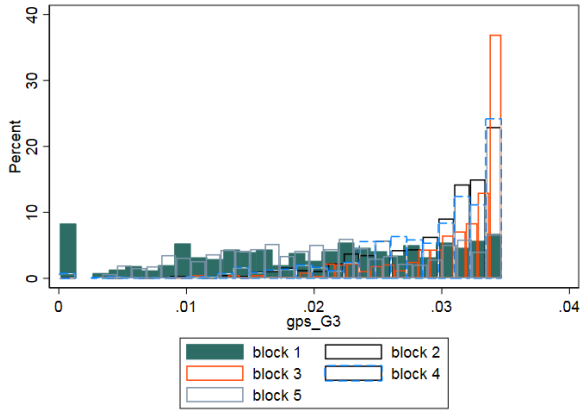
(f) GPS at p30 - after step 3



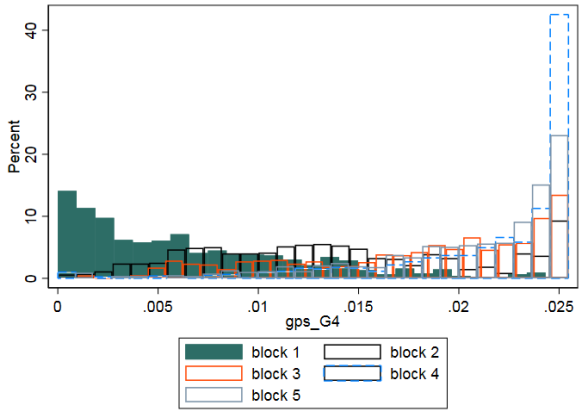
(g) GPS at p50 - before CS condition



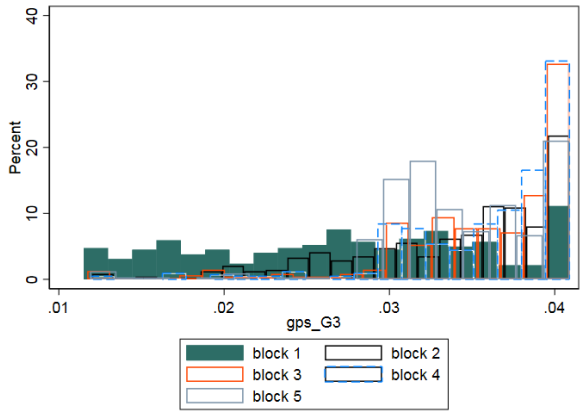
(h) GPS at p70 - before CS condition



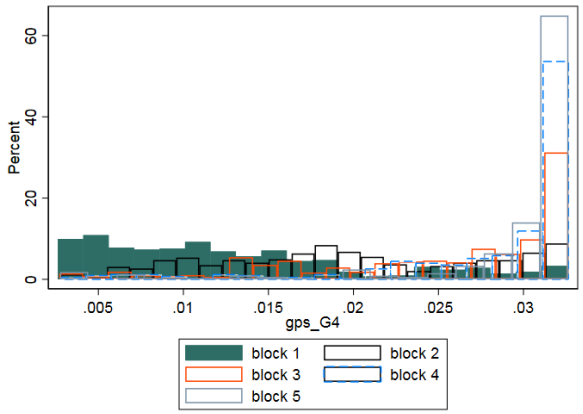
(i) GPS at p50 - after step 1



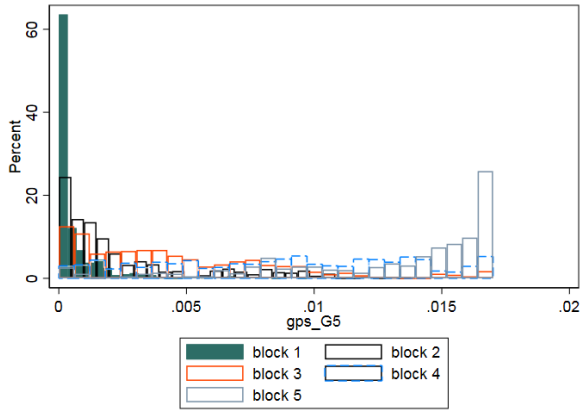
(j) GPS at p70 - after step 1



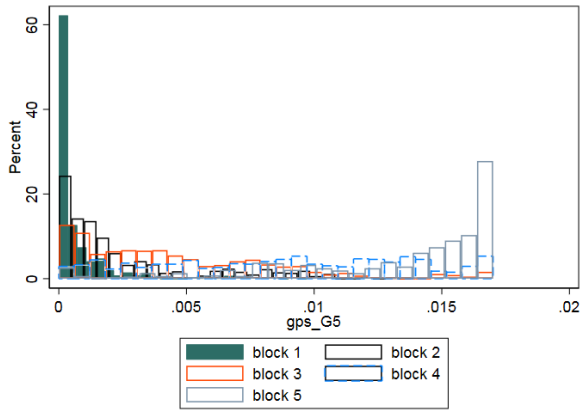
(k) GPS at p50 - after step 3



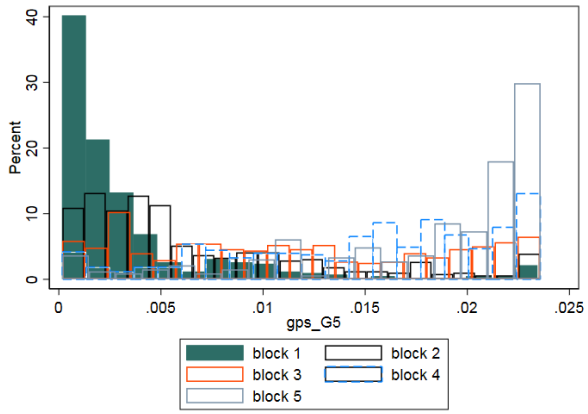
(l) GPS at p70 - after step 3



(m) GPS at p90 - before CS condition



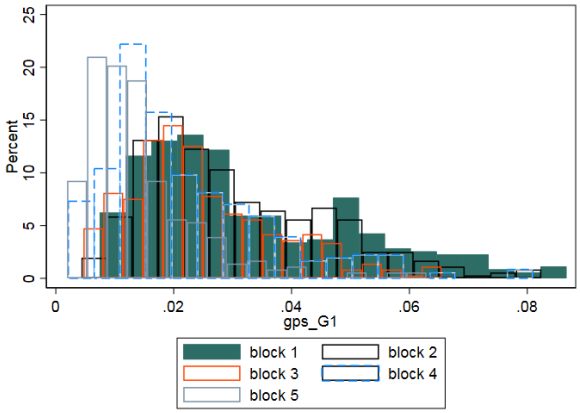
(n) GPS at p90 - after step 1



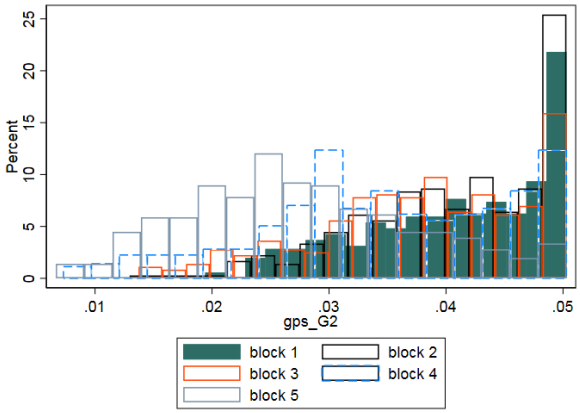
(o) GPS at p90 - after step 3

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 4.2, own calculations.

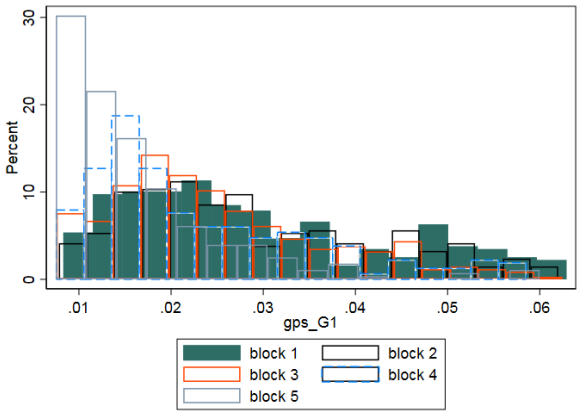
Figure 4.E.6: Common support - A2 sample - small municipalities



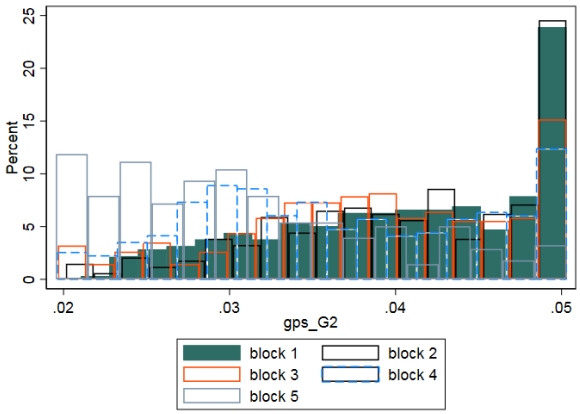
(a) GPS at p10 - before CS condition



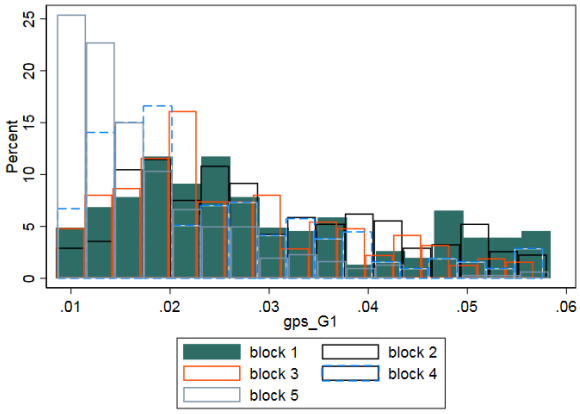
(b) GPS at p30 - before CS condition



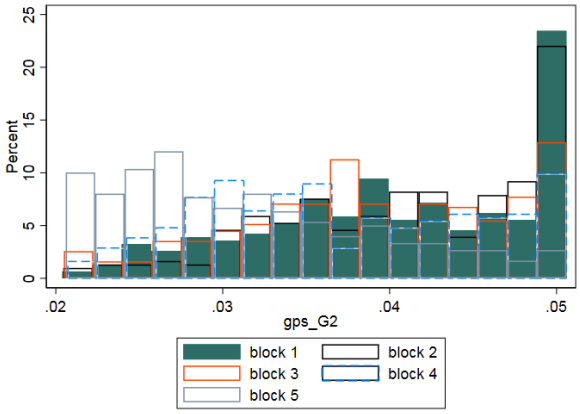
(c) GPS at p10 - after step 1



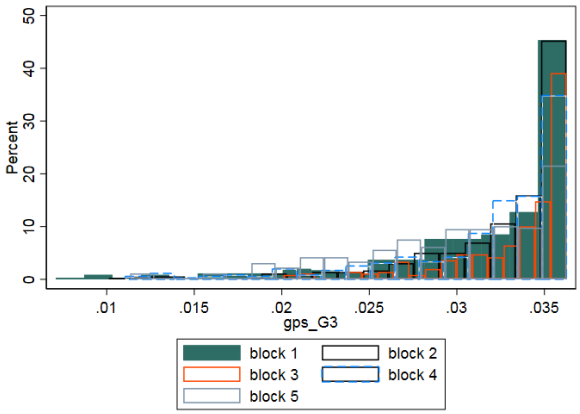
(d) GPS at p30 - after step 1



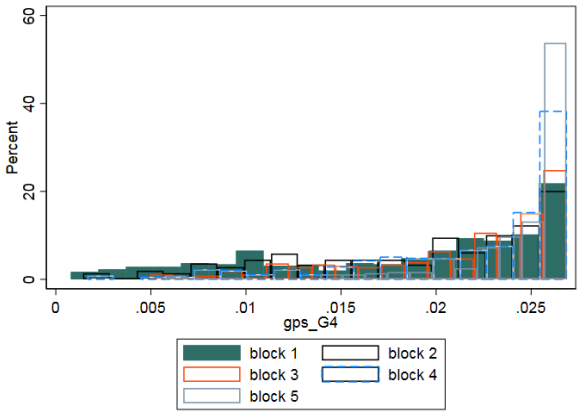
(e) GPS at p10 - after step 3



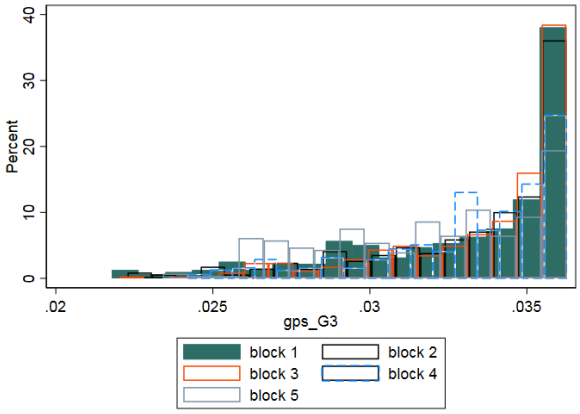
(f) GPS at p30 - after step 3



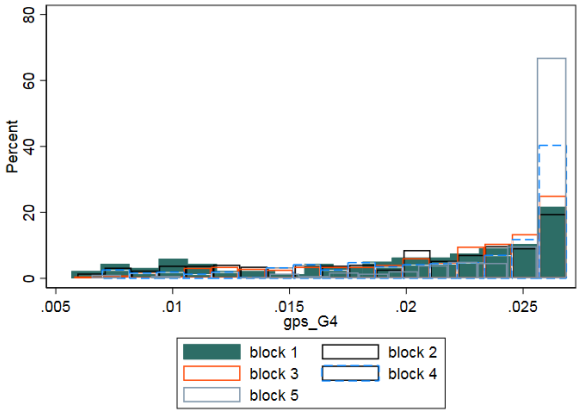
(g) GPS at p50 - before CS condition



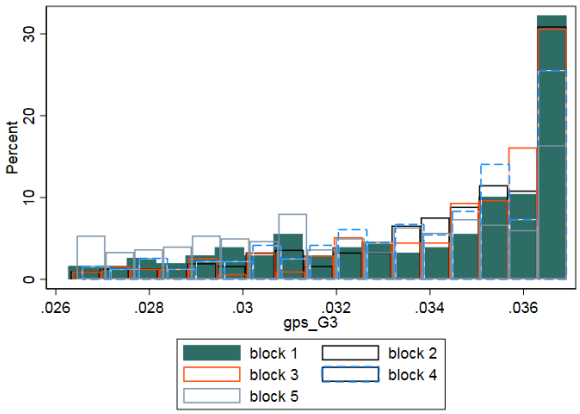
(h) GPS at p70 - before CS condition



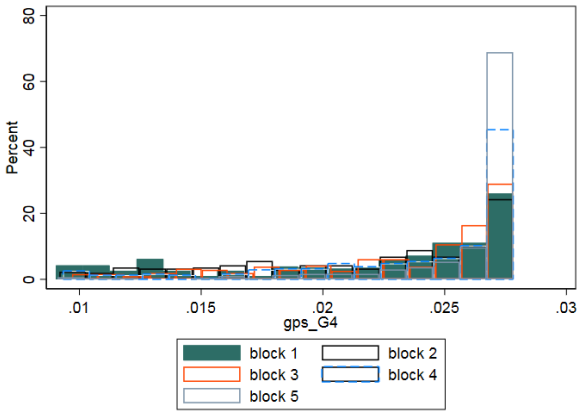
(i) GPS at p50 - after step 1



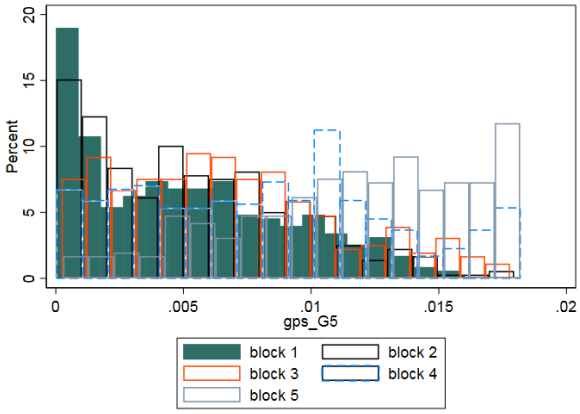
(j) GPS at p70 - after step 1



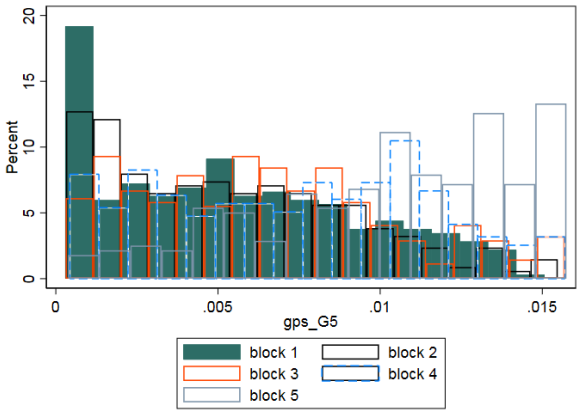
(k) GPS at p50 - after step 3



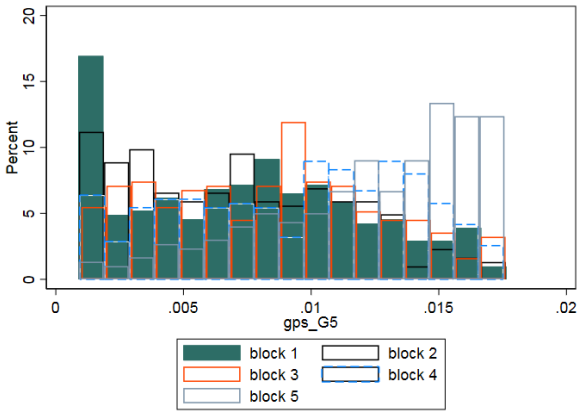
(l) GPS at p70 - after step 3



(m) GPS at p90 - before CS condition



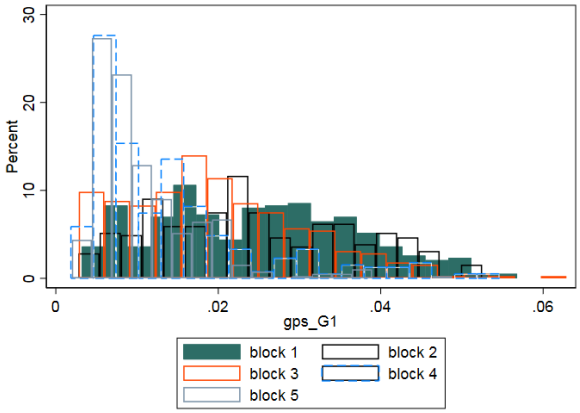
(n) GPS at p90 - after step 1



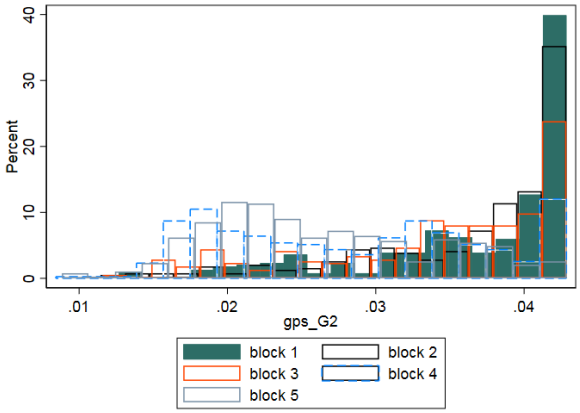
(o) GPS at p90 - after step 3

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 4.2, own calculations.

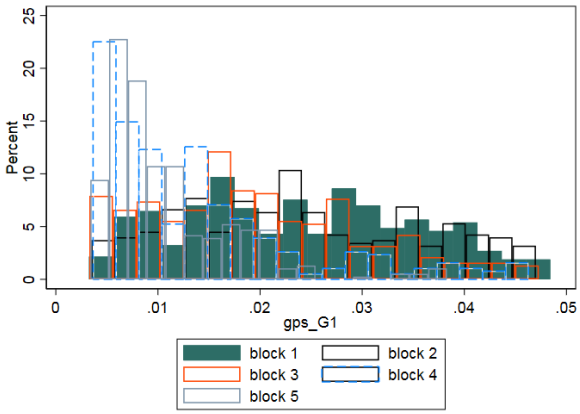
Figure 4.E.7: Common support - A2 sample - large municipalities



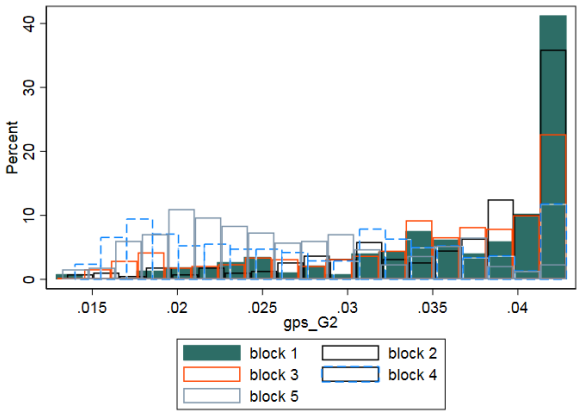
(a) GPS at p10 - before CS condition



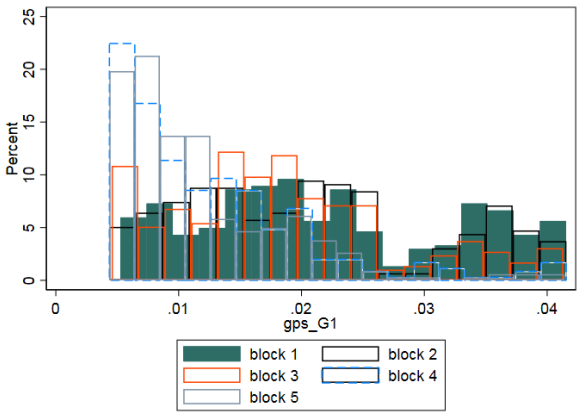
(b) GPS at p30 - before CS condition



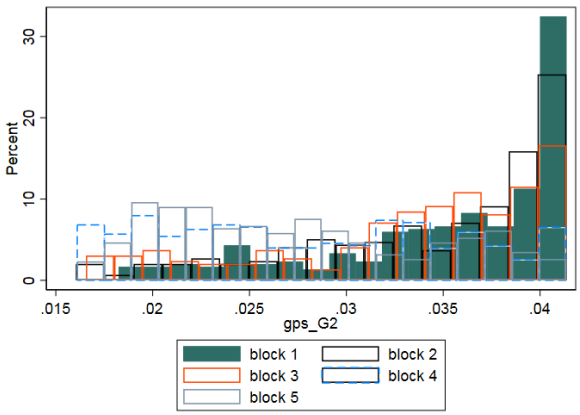
(c) GPS at p10 - after step 1



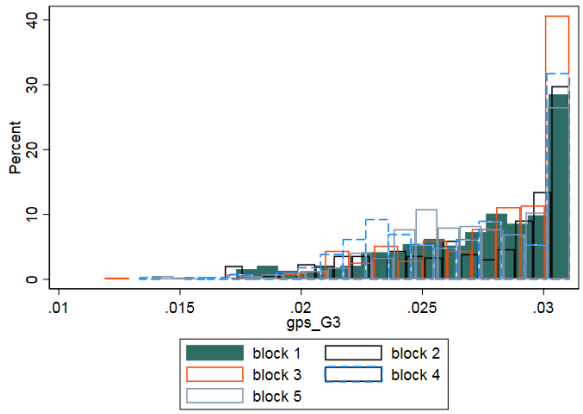
(d) GPS at p30 - after step 1



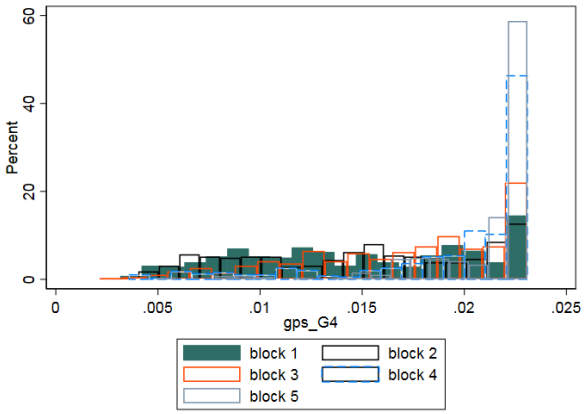
(e) GPS at p10 - after step 3



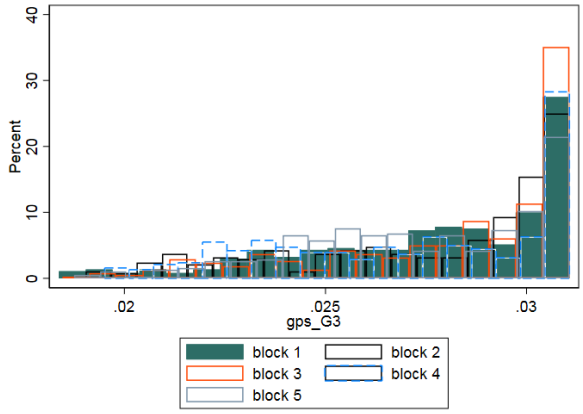
(f) GPS at p30 - after step 3



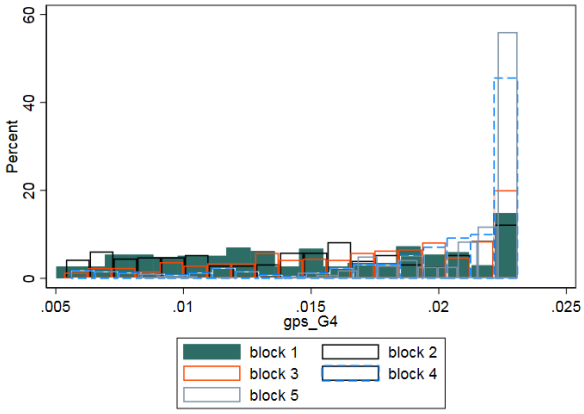
(g) GPS at p50 - before CS condition



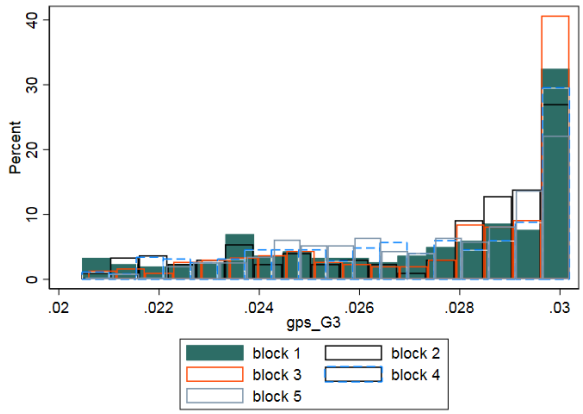
(h) GPS at p70 - before CS condition



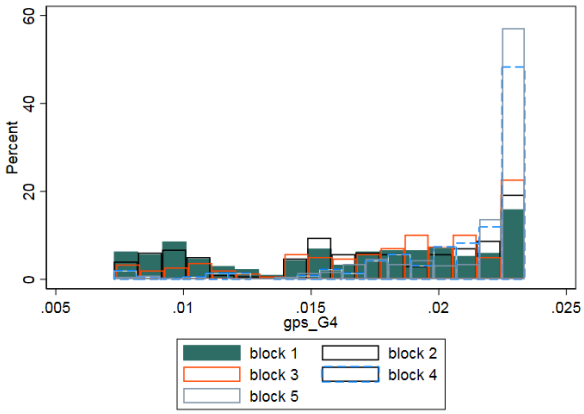
(i) GPS at p50 - after step 1



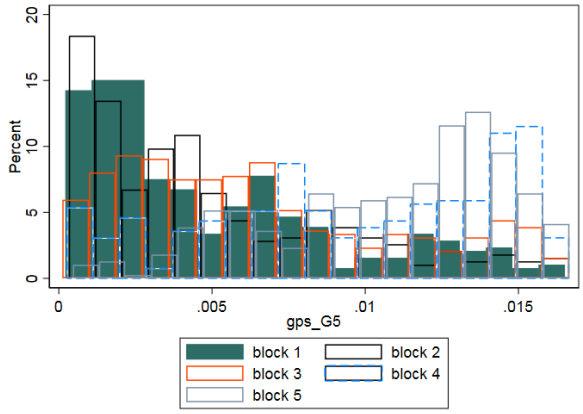
(j) GPS at p70 - after step 1



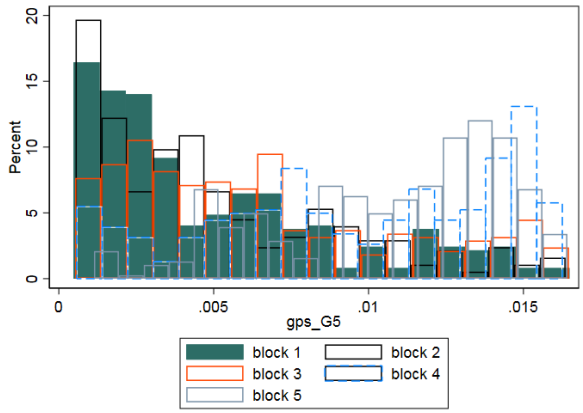
(k) GPS at p50 - after step 3



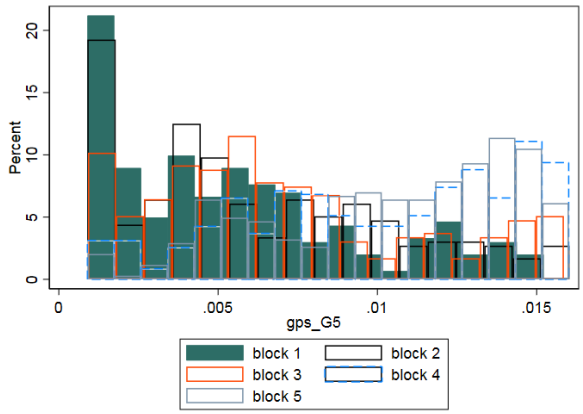
(l) GPS at p70 - after step 3



(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 4.2, own calculations.

Table 4.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	SD	Mean 1	Mean 3	SD	Mean 1	Mean 4	SD	Mean 1	Mean 5	SD	Mean 2	Mean 3	SD
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	SD	Mean 2	Mean 5	SD	Mean 3	Mean 4	SD	Mean 3	Mean 5	SD	Mean 4	Mean 5	SD
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 4.2, own calculations.

Table 4.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 4.2, own calculations.

Table 4.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	SD	Mean 1	Mean 3	SD	Mean 1	Mean 4	SD	Mean 1	Mean 5	SD	Mean 2	Mean 3	SD
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	SD	Mean 2	Mean 5	SD	Mean 3	Mean 4	SD	Mean 3	Mean 5	SD	Mean 4	Mean 5	SD
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 4.2, own calculations.

Table 4.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 4.2, own calculations.

Table 4.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 4.2, own calculations.

Table 4.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 4.2, own calculations.

Table 4.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 4.2, own calculations.

Table 4.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	SD	Mean 1	Mean 3	SD	Mean 1	Mean 4	SD	Mean 1	Mean 5	SD	Mean 2	Mean 3	SD
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	SD	Mean 2	Mean 5	SD	Mean 3	Mean 4	SD	Mean 3	Mean 5	SD	Mean 4	Mean 5	SD
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 4.2, own calculations.

Table 4.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 4.2, own calculations.

Table 4.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	SD	Mean 1	Mean 3	SD	Mean 1	Mean 4	SD	Mean 1	Mean 5	SD	Mean 2	Mean 3	SD
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	SD	Mean 2	Mean 5	SD	Mean 3	Mean 4	SD	Mean 3	Mean 5	SD	Mean 4	Mean 5	SD
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

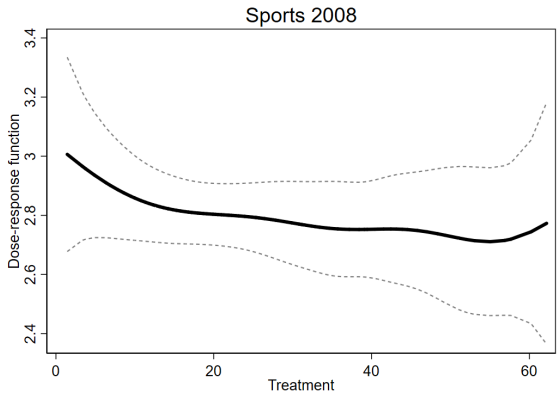
Sources: RDC of the Federal Statistical Office and Statistical Offices of the Länder (2018a; 2018b), SOEP as well as further data sources as discussed in Chapter 4.2, own calculations.

Table 4.E.18: Balancing after step 2 of common support procedure - A2 sample - large municipalities

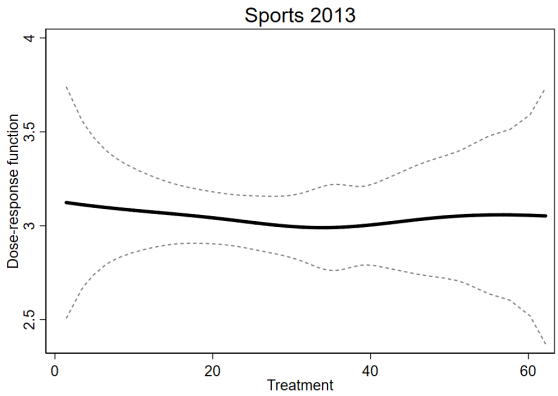
	Block 1 vs 2	Block 1 vs 3	Block 1 vs 4	Block 1 vs 5	Block 2 vs 3	Block 2 vs 4	Block 2 vs 5	Block 3 vs 4	Block 3 vs 5	Block 4 vs 5
ln(Population per km2/100)	15.91	8.67	-24.01	3.32	-7.68	-42.17	-14.46	-34.52	-6.25	30.50
ln(Average rain/100)	-10.32	-3.13	34.85	7.79	7.85	45.93	18.65	40.34	11.60	-28.22
ln(Proportion 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 4.2, own calculations.

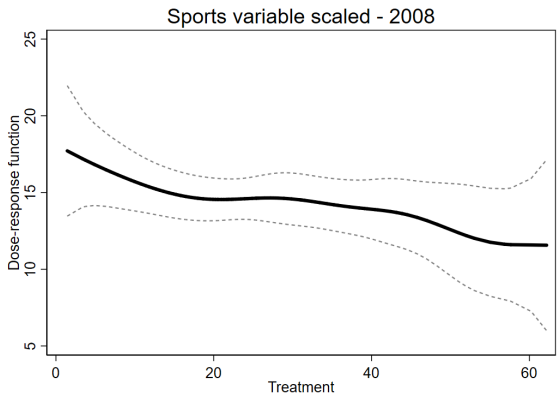
Figure 4.E.8: Additional results A1 sample (17-30 years of age)



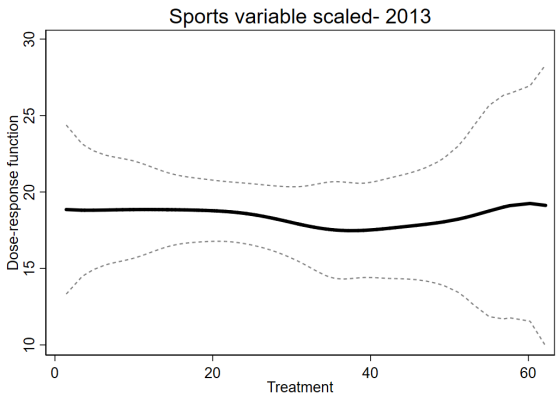
(a) DRF sports 2008



(b) DRF sports 2013



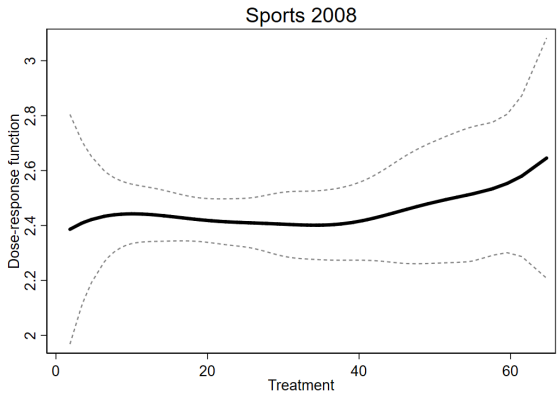
(c) DRF sports variable scaled - 2008



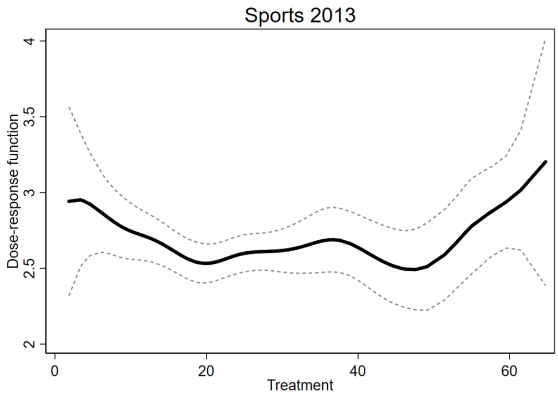
(d) DRF sports variable scaled - 2013

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 4.2, own calculations.

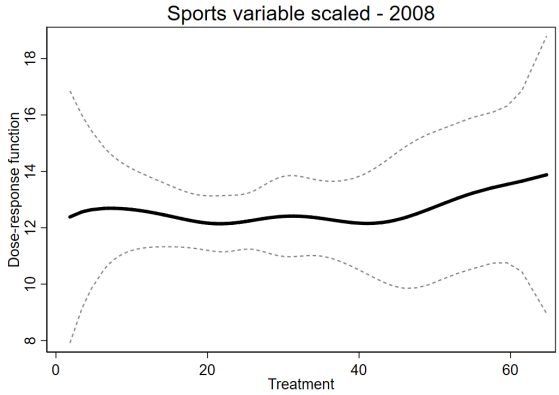
Figure 4.E.9: Additional results A3 sample (46-65 years of age)



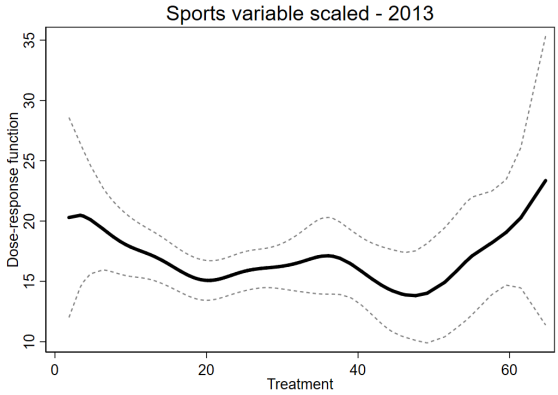
(a) DRF sports 2008



(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 4.2, own calculations.

5 Individual labor market effects of local public expenditures on sports

with Tim Pawlowski, Tim Wallrafen and Michael Lechner[1]

Abstract

By merging administrative data on public finances of all municipalities in Germany with individual data coming from the German Socio-Economic Panel, we explore whether local public expenditures on sports facilities may influence individual labor market outcomes. Our identification strategy follows a selection-on-observables approach and exploits the panel structure of the data covering 12 years between 2001 and 2012. The results of our matching estimations suggest that both women and men living in municipalities with high expenditure levels benefit with approximately 7 percent of additional household net income on average. This effect, however, is fully captured by earning gains for men rather than women living in the household. Additional analysis suggests, that these gender differences, which can also be observed in terms of working time, hourly wage and employment status, appear plausible since women in the age cohort under consideration are less likely than men to engage in sports in general and in any of the publicly funded sports facilities in particular. Moreover, improved well-being and health are possible mechanisms that determine how the positive labor market effects for men may unfold.

Keywords: Labor market effects, public expenditures, sports, health, well-being

JEL codes: Z20, H72, H75, J31

[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 Ute Schüttoff for her help in collecting, cleaning and transforming the regional data. A previous version of the paper was presented at the 88th Annual Meeting of the Southern Economic Association in Washington and in the Research Seminar on Sports Economics at the University of Tübingen. We thank participants for helpful comments and suggestions. The usual disclaimer applies.

5.1 Introduction

Although the literature analyzing the effects associated with public expenditures on health (e.g., Crémieux, Ouellette, & Pilon, 1999), education (e.g., Jackson, Johnson, & Persico, 2016) and labor market programs (e.g., Ham, Swenson, İmrohoroğlu, & Song, 2011) is rich, little is known about related effects of public expenditures in other areas such as sports. This is surprising, given (i) the evidence, that sports participation might improve health (e.g., Charness & Gneezy, 2009), educational (e.g., Stevenson, 2010) and labor market outcomes of individuals (e.g., Lechner, 2009) and (ii) the fact, that policy makers commonly legitimate public expenditures on sports with the objective to foster such effects (e.g., Federal Ministry of Interior, 2010).

We use nation-wide micro-level panel data to explore whether sports related public expenditures at the municipality level (SPE) may unfold any beneficial effects at the individual level. In this regard, our study is both conceptually motivated by and closely related to the study by Jackson et al. (2016), who explore the effects of public district school spending on educational and labor market outcomes of adults living in the USA. Our setting is Germany, where the total amount of annual sports related public expenditures is estimated at around €10 billion. Hereof, the largest shares (i.e. about €3-4 billion each) are spent on physical education (P.E.) teachers in schools as well as the construction and maintenance of sports infrastructure (Pawlowski & Breuer, 2012). We focus on the latter since sports infrastructure forms the basis for sports activity across all age groups and recently also became an explicit policy target in order to reduce physical inactivity in society.[2] Since money spent by regional authorities on sports facilities is commonly channeled through local authorities, we focus on the lowest administrative level, that is, municipalities.

The expenditure data come from accounting records of the more than 12,000 municipalities in Germany and are matched with meteorological data, information on the socio-economic characteristics of the local population as well as extensive individual data

[2]For instance, the World Health Organization (WHO) recommends amongst others (World Health Organization, 2013, p. 34): *"Creation and preservation of built and natural environments which support physical activity in schools, universities, work-places, 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."*

coming from a representative sample of German households, the German Socio-Economic Panel (SOEP). To consider that individuals may not only benefit from the expenditures of their own municipalities but also neighboring ones, we construct a distance weighted expenditure measure.

Because public expenditures are policy variables and individuals may deliberately make location choices based on the availability of sports infrastructure, we follow a selection-on-observables approach and exploit the panel structure of our data in order to identify the effects of interest. In this regard, we control for regional characteristics (such as population size, population density, weather conditions and geographical information) and individual characteristics (such as educational level, family status and household composition) expected to jointly influence public expenditures and labor market outcomes. Moreover, by conditioning on pre-exposure sports participation, labor market performance and different health and social capital indicators we are able to control to some extent for unobservable confounding factors.

Overall, our matching analysis reveals sizable effects of local sports facility expenditures on individual labor market outcomes. For both men and women, we observe a significant (average) increase in household monthly net income of about €190, or 7 percent, when moving from medium to high levels of SPE. Interestingly, however, these effects are fully captured by earnings gains for men rather than women living in the household.[3] Such gender differences can also be observed in terms of working time, hourly wage and employments status. While we do not observe any statistically and economically significant effects for women, we find that (on average) men work an hour more and benefit by a €1 increase in hourly wage as well as an increase in the probability of being employed full time when moving to comparably higher levels of SPE. Finally, we find that older rather than younger men benefit from comparably higher SPE levels.

Additional analysis based on the SOEP-Innovation Sample (SOEP-IS) suggests that the revealed effect differences appear plausible.[4] For example, women in the age cohort under consideration are less likely than men to engage in sports *in general* and use any of the publicly funded sports facilities *in particular*. Moreover, younger men, though they

[3]In general, this finding is reinforced by our results for the comparison between low and high SPE levels. However, the estimated effects for household income are less precise and not significant.

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

are more active than older men, have a comparably higher probability of engaging in sports outside any of the publicly funded sports facilities. Finally, while possibilities to test whether these effects unfold due to (SPE induced) differences in sport participation patterns are limited, we do find some evidence that improved well-being and health are possible mechanisms that determine how these effects may unfold.

The structure of the paper is as follows. Section 5.2 presents the institutional setting and the sports related expenditure data. Section 5.3 outlines the econometric specification, including details about the sampling, the identification strategy as well as the estimation scheme. Our findings are discussed in Section 5.4. Section 5.5 concludes. Appendix 5.A offers more information on the data cleaning, transforming, and enriching process. Appendix 5.B provides a description of variables and further estimation results including information about the operational characteristics of the estimators. Appendix 5.C contains results of the robustness checks. Appendix 5.D displays the results about effect heterogeneity and Appendix 5.E reports the results about possible mechanisms that determine how the effects may unfold.

5.2 Sports related public expenditures: data and measure

Sports in Germany are subsidized at all three administrative levels, i.e. by the federal, regional and local authorities.[5] Regional and local authorities in particular aim to foster leisure sports participation (Pawlowski & Breuer, 2012). Since money spent on sports facilities by regional authorities is commonly channeled through local authorities, we focus our analysis on the lowest administrative level: district and urban municipalities/cities including Hamburg, Bremen, and Berlin.[6]

Expenditure data from these more than 12,000 municipalities come from the Research Data Center (RDC) of the RDC of the Federal Statistical Office and Statistical Offices of the Länder (2018b) and cover the financial years between 2001 and 2006. They include

[5]Note, Steckenleiter, Lechner, Pawlowski, and Schütoff (2019) use the same data analyzing the impact of SPE on the probability to engage in sports (see the discussion in Section 5.4.5). Therefore, Chapter 5.2 has certain overlaps with the data description in Steckenleiter et al. (2019). However, including this information is indispensable to fully shed light on the empirical strategies employed in both papers.
[6]These three cities are independent from any of the other 13 states ('Länder') in Germany.

expenditures for the construction and maintenance of sports facilities like leisure-time sports facilities, sports airfields, sports stadiums, sports fields, tennis courts, toboggan and bobsled runs, sports schools, gyms, ski jumps, indoor or outdoor pools as well as Olympic sports facilities (Federal Statistical Office, 2016).[7] Since individuals may not only benefit from the expenditures (and the resulting services and facilities) of their own municipalities, but also of neighboring ones (though to a lesser extent due to higher travel costs), we construct a distance weighted expenditure measure instead of taking the expenditure from the municipality of residence. Our measure considers for an individual k living in municipality m per-capita expenditures of the municipality of residence D_m as well as the sports related expenditures of all neighboring municipalities D_i within a certain radius around the municipality of residence:

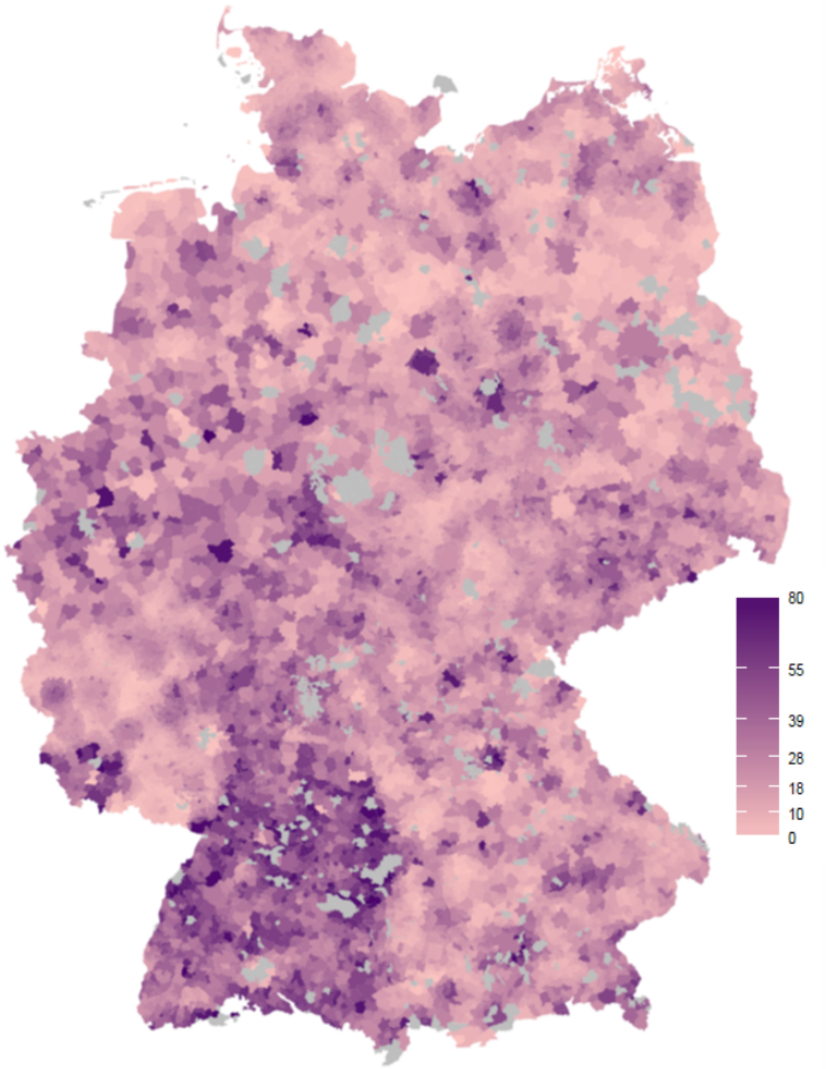
$$D_k = \frac{x}{r} D_m + \left\{ \frac{(r-x)}{r} \left[\frac{\sum w_i^{'} D_i}{\sum w_i^{'} p_i} \right] \right\} \tag{5.1}$$

with x measuring the distance between the geometric centers (calculated as polygon centroids) of the closest neighboring municipality and the municipality of residence, r measuring the radius, p measuring the population size of the neighboring municipalities, and $w_i^{'} = w_i / \sum_{j=1}^{N} w_j$, $w_i = 1/d_i$ as further weight with d measuring the distance to the geometric center. In summary, this measure ensures, that (i) the shares of expenditures for all municipalities under consideration for constructing each D_k add up to 100 percent and that the expenditures of the municipality of residence weights more (ii) the larger the municipality of residence is and (iii) the further the neighboring municipalities are away.

Figure 5.1 shows the distribution of these distance weighted per capita expenditures (five-year-average, 2002-2006) across Germany. Since some outliers are present, we kept

[7] To work with these data, all figures were converted to 2004 Euros using the Consumer Price Index from the Federal Statistical Office (2017) as well as cleaned, transformed and enriched in several steps. For instance, since an increasing number of municipalities have transformed sports facilities (in particular swimming pools) into owner-operated community enterprises (OCPs) over time, we added financial data for all sports related OCPs coming from the annual balance sheets of public funds, institutions and enterprises in Germany (RDC of the Federal Statistical Office and Statistical Offices of the Länder, 2018a). Moreover, to consider the fact that some funded sports facilities (in particular soccer stadiums) are predominantly used by professional sports clubs only, we reduced all figures by income from rental fees. The intuition behind this approach is that comparably larger construction and maintenance costs for sports facilities such as professional soccer stadiums come along with comparably higher rental fees. In line with this, we reduced all figures in the twelve cities hosting any of the FIFA World Cup 2006 games by the amounts spent for the construction and renovation of stadiums for this event. More information on the data cleaning, transforming and enriching process is provided in Appendix 5.A.

Figure 5.1: Distance weighted sports facility per capita expenditures in Germany.



Notes: This figure displays mean (2002–2006) net values (i.e., based on expenditure grouping number 999 minus income grouping number 399) in 2004 euros between 0 and the 99th percentile (gray areas display municipalities with corresponding expenditure levels outside this range). The underlying shapefile for this map is provided by ©GeoBasis-DE/BKG 2017. Data sources: RDC of the Federal Statistical Office and Statistical Offices of the Länder (2018a, 2018b), own calculations.

observations with positive SPE values only up to the 99th percentile in this figure as well as in our estimations.[8] While comparably larger amounts are spent in the state of

[8]By construction, our expenditure measure has negative values for municipalities displaying sports facility related income (i.e. rental fees) higher than the corresponding (running) costs. Moreover, some municipalities spent considerably more than the rest during the observation period. We decided to delete these municipalities from our analysis (denoted as gray areas in Figure 5.1). Note, while Figure 5.1 is based on all municipalities in Germany, the estimation sample only consists of the subsample of municipalities matching any of the SOEP observations. Therefore, some deviations occur with regard to the percentile figures. In this regard, SPE values until the 99th percentile in our estimation sample measure up to €85 (instead of €80 in the full sample, see Figure 5.1).

Baden-Wuerttemberg located in south-west Germany, a significant within-state variation in SPE is generally evident across Germany, underscoring the relevance of analyzing the lowest administrative level in our study as argued by Harris, Evans, and Schwab (2001).

5.3 Empirical strategy

We explore the effects of local SPEs from 2002 to 2006 on labor market outcomes in the period 2007 to 2012 by following and comparing individuals with similar characteristics in 2001 over time. This is done by merging our distance weighted expenditure measure as defined in Section 5.2 with individual information from the German Socio-Economic Panel (SOEP)[9] on a yearly basis. Merging on a yearly basis allows to allocate 'true' expenditure values also to individuals who relocated in the exposure period 2002-2006 before averaging the expenditure measure over five years. Averaging the expenditure measure seems reasonable to smooth out any sports facility 'investment shocks' that might occur over time. To work with this panel data, we need to impose a couple of sample restrictions as explained in Section 5.3.1. Moreover, to tackle selection and endogeneity issues, we exploit the panel structure of our data and apply a selection-on-observables approach by controlling for relevant variables jointly influencing local sports expenditures and our outcome variables. These confounders include several regional characteristics (such as population size and density as well as weather conditions and geographical information) as well as individual characteristics (such as socio-demographics, information on sports participation behavior as well as different education, health, social capital and labor market indicators) which are taken from the pre-exposure year 2001 (see Section 5.3.2). To estimate the effects of interest, we discretize our expenditure measure in three states by cutting at the 33rd and 66th percentile and employ radius matching with bias adjustment (see Section 5.3.3).

[9]The SOEP is a longitudinal panel survey containing a representative sample of German households since 1984. Data are merged by using the official residence municipality key code ('Amtlicher Gemeindeschlüssel – AGS'). Since SOEP data including the official residence municipality key code is subject to severe data restriction policies, data merging and analysis took place in the Research Data Center of the SOEP in Berlin.

5.3.1 Sampling

As explained previously, our panel data cover an observation window of 12 years: the pre-exposure year 2001, an exposure period between 2002 and 2006 and a post-exposure period between 2007 and 2012, during which we estimate the labor market effects. Similar to Lechner (2009) or Lechner and Sari (2015), we are predominantly interested in long-run labor market outcomes. A long observation window, however, comes along with two problems, an greater panel attrition as well as an increasing probability of retirement for older individuals.[10] Therefore, we decide to restrict our analysis to a six-years (post-exposure) observation window between 2007 and 2012. To work with this panel, we first restrict our sample to individuals who have a valid interview in 2001 and the first outcome year under consideration (i.e. 2007). At this stage, we also exclude individuals with missing values in either lagged expenditure, lagged outcomes or any confounding variable used in our models (see Section 5.3.2). Second, we restrict our sample to all individuals aged 20 and older in the pre-exposure year and also excluded individuals not yet graduated in 2001. Third, we impose an upper age limit of 52 years in 2001 since these individuals are 63 in 2012 and as such still below the average age of retirement (German Statutory Pension Insurance Scheme, 2017). Moreover, we exclude any individuals who were still in the education system in 2007 onwards (e.g. trainees, students, civilian servants, conscripts). Finally, to avoid the results being driven by outliers in expenditures we restrict the sample to observations up to the 99th percentile and only kept observations with positive values for the expenditure variable. These restrictions resulted in a data set with 3,071 men and 3,427 women for our analysis in 2007. To avoid further efficiency loss in the first-year models, we opted against a balanced panel approach, which would have required that no observation have any missing data in any of the other outcome years.

[10] While we are able to test for any panel attrition in our study (see Section 5.4.3), smaller samples generally suffer from efficiency loss.

5.3.2 Identification

Since public expenditures are policy variables[11] and individuals may deliberately make location choices based on the availability of sports infrastructure[12], selection and endogeneity issues are obvious concerns in our setting. To tackle these concerns, we exploit the panel structure of our data and apply a selection-on-observables approach. The key idea of this approach is, that after conditioning on observed confounding variables, selection into the level of SPE is random (Imbens, 2004). To ensure that this so-called conditional independence assumption (CIA) holds, we need to control for all variables jointly influencing local public expenditures on sports facilities and individual labor market outcomes. These confounders include various regional and individual characteristics and will be explained in the following.[13]

Our first set of variables covers several regional characteristics. These variables are selected to control for any confounding influences at the aggregate (municipality) level. (i) In this regard, we consider population size as a relevant predictor for both, the availability and conditions of local (sports) infrastructure[14] as well as the labor market conditions and prospects of individuals. To better consider the latter aspect, we use population size in the corresponding BIK region instead of population size in the municipality only.[15] Similar arguments hold with regard to the population density, i.e. rural vs. urban areas. Since population density is not directly considered in the BIK regions, we include (ii) population per square kilometer and also take (iii) its interaction with the different BIK regions into account. To directly control for the economic conditions as well as the level of education within the municipalities, we also include (iv) the share of unemployed in the working age population as well as (v) the share of graduates with higher education

[11] For instance, local authorities may use observed sports activity levels in order to determine SPE levels or they may be influenced by neighboring municipalities' spending behaviors, i.e. spatial dependency between local authorities as discussed for instance by Moscone, Knapp, and Tosetti (2007).

[12] This behavior can be described as a kind of Tiebout sorting effect (Tiebout, 1956), that is, the consumer-voter moves to that municipality whose local government best satisfies his or her set of preferences.

[13] All confounders are measured in the pre-exposure year (2001) if not stated otherwise.

[14] We expect that bigger municipalities have generally better opportunities to build and maintain such infrastructure due to economies of scale.

[15] A BIK region (BIK–Aschpurwis & Behrens GmbH, 2001) is made up by all municipalities that have at least 7 percent of their employees (subject to social insurance contributions) commuting to the same core city. We take BIK regions with 500,000 and more inhabitants as a reference category and differentiate between three further categories as dummies in our models, i.e. BIK regions with 0-49,999, 50,000-99,999 and 100,000-499,999 inhabitants.

entrance qualification of total graduates.[16] While it is common practice in the literature on public spending to control for socio-cultural conditions in the municipalities in general (e.g., Busemeyer, 2007), it appears to be particularly relevant to control for the composition of the population with migration background in our setting given the popular policy claim that sport might serve as a tool for integration (e.g., German Olympic Sports Confederation, 2019). Although we are not able to implement any fine measure here, we control for (vi) the share of foreigners in a municipality. Weather conditions are also expected to influence the investment and maintenance costs for (sports) infrastructure in a municipality as well as the labor market outcomes of individuals. While the effects of temperature, precipitation and windstorms on labor productivity and other economic outcomes are already well documented in the literature (see Dell, Jones, and Olken (2014) for a review), this is to the best of our knowledge not the case with regard to possible effects on infrastructure investments and maintenance costs. However, it appears plausible to assume, that practicing outdoor sports is less popular in rainy areas, while providing opportunities for outdoor sports such as public parks or green areas is probably less cost-intensive than providing and maintaining indoor sports facilities. Therefore, we also control for (vii) the average sum of precipitation as well as (viii) the average sunshine duration per year.[17] Finally, we control for (ix) the region where the municipality is located, i.e. south (including the states Baden-Wuerttemberg and Bavaria), north (including the states Lower-Saxony and Schleswig-Holstein as well as the city states Bremen and Hamburg), west (including the states of Hessen, Rhineland-Palatine and Saarland) and east (including the states of Brandenburg, Mecklenburg-West Pomerania, Saxony, Saxony-Anhalt and Thuringia as well as the city state Berlin). Controlling for the region in Germany is important given the significant regional variation in sports facility spending with comparably higher expenditures in southern Germany (see Figure 5.1) which is also a comparably more prosperous region with better labor market conditions and prospects for individuals.

Our second set of variables covers several individual characteristics coming from the

[16] For both measures, we take the six-years average (2001–2006) to avoid picking up any year-specific fluctuation.

[17] Similar to the share of unemployed in the working age population and the share of graduates with higher education entrance qualification of total graduates, we intend to avoid picking up any year-specific fluctuations. Since, weather conditions are subject to changes both in the short- and long-term we decided to consider long-term averages covering the period 1981 – 2010.

SOEP. While the regional characteristics are used in order to control for any confounding influences at the aggregate (municipality) level, the individual characteristics condition on top on factors that are expected to influence both the labor market performance and the level of SPE at the disaggregated (individual) level. In this regard, we do not expect any individuals (other than the decision makers in the local government) to directly influence the level of SPE. However, (sports active) individuals may deliberately make location choices based on the availability and quality of local sports infrastructure.[18]

To systematically approach this potential confounding issue, we control for all characteristics that were previously found to jointly influence sports participation behavior and labor market performance of individuals (for an overview, see Lechner, 2009). These variables include (i) being married or (ii) being divorced, (iii) having German citizenship, (iv) number of children in the household and (v) age. In addition, we include two sets of educational measures in our models. The first set considers the (vi) highest school degree (measured as lower secondary school or no school degree, intermediate secondary school degree and upper secondary school degree or other graduation diploma). The second set considers (vii) individuals' level of (higher) education/training (measured as still pursuing education, no vocational degree, vocational degree or university degree). To directly control for differences in sports participation behavior in the pre-exposure period, we also include (viii) the available indicator in the SOEP measuring whether an individual participated at least monthly in sports. Moreover, we also consider differences in labor market performance in the pre-exposure period directly by including (ix) monthly household net income, (x) gross earnings, (xi) working time per week, (xii) hourly wage as well as (xiii) full-time employment.[19] Finally, we also control for different health and social capital indicators from the pre-exposure period since both might influence sports participation behavior and labor market performance in the subsequent years. In this regard, we control

[18] While we do not know anything about the extent to which leisure-time sports facilities might stimulate location choices in practice, there is some evidence that income and family status might increase the probability of living closer to professional sports facilities (e.g. Ahlfeldt & Kavetsos, 2014). Moreover, several studies previously looked at the socio-economic determinants of location choice in general. In this regard, there is some evidence for the US that comparably less populated areas have relatively higher marriage rates (Lichter, LeClere, & McLaughlin, 1991) while the probability to live in suburban areas increases with increasing age (Alba & Logan, 1991).

[19] All monetary measures used in our estimations were converted to 2004 Euros using the Consumer Price Index from the Federal Statistical Office (2017). Hourly wage is not readily available in the SOEP and was computed by dividing monthly gross earnings by weekly hours (x 4.3). We decided to also control for hourly wage in our estimations in order to allow for more flexibility.

for (xiv) satisfaction with health, (xv) satisfaction with life, (xvi) the number of doctor visits per year as well as (xvii) the physical component (PCS) and (xviii) mental component summary scales (MCS).[20] The available three social capital indicators are dummies measuring interpersonal networks, i.e. (xix) whether an individual was at least monthly helping out friends, neighbors or relatives, and social engagement by (xx) the frequency of volunteering work in clubs or social services (at least monthly) as well as (xxi) the general involvement in a citizens' group, political party and/or local government. This rich set of regional and individual characteristics allows us to control for the most relevant observable confounders. Moreover, while we are unable to control for genetics or any psychological, cognitive or emotional factor which might influence both sports participation behavior and individual labor market performance directly, we argue that conditioning on pre-exposure sports participation, labor market performance and the different health and social capital indicators enables us to indirectly control at least to some extent for such unobservable confounders. In this regard, the available measures 'control' for some sort of individual fixed effects as argued by Lechner (2009).

5.3.3 Estimation

We opt for a matching type estimator since it is more flexible and more robust with respect to the statistical assumptions imposed (Imbens, 2004). Similar to Lechner and Sari (2015), who estimated the labor market effects for three different sports activity levels (active versus moderately active versus inactive), we discretize our SPE measure as defined in Section 5.2 in three states: low (€0-€20), medium (€20-€31) and high (€31-€85), using the 33rd and 66th percentiles as cutoffs. Doing so allows us to compute and compare overall three average treatment effects (ATEs): the labor market effects of moving from low to medium, from low to high as well as from medium to high SPE levels. Since the CIA is assumed to hold across all strata, the sample reduction results of Lechner (2001) apply, that is, for estimating either of these three effects, participants in the other treatment state are deleted for the purpose of this particular estimation. This,

[20]The latter two measures are superordinate scales available since 2002 on a biannual basis and had been derived by Andersen, Mühlbacher, Nübling, Schupp, and Wagner (2007) from an internationally applied inventory of health measures (the so called SF-12v2 indicators) using explorative factor analysis. Since the physical component (PCS) and mental component summary scales (MCS) are not available in 2001, we took both measures from the 2002 SOEP survey.

however, requires that observations on support for a certain group are the same across estimations. In other words: when estimating the effects of low versus high and medium versus high levels of SPE, the same observations must be considered in the treatment group (high SPE level). This is ensured by implementing an iterative approach, that is, we deleted any observations off support after each estimation and re-estimated all models (including propensity scores) until all remaining observations were on support.[21] Finally, due to considerable gender specific differences in terms of both sports activity patterns (e.g., Eccles & Harold, 1991) and labor market participation (e.g., Fitzenberger, Schnabel, & Wunderlich, 2004), we decided to estimate all effects for women and men separately.[22] We estimate all effects by radius matching with bias adjustment as Lechner, Miquel, and Wunsch (2011) suggest. A simulation study (Huber, Lechner, & Wunsch, 2013) shows that this matching estimator performs particularly well.[23] Inference is based on boot-strapping the sample 99 times.[24]

5.4 Results

We first describe our findings for the selection process (Section 5.4.1). We then turn to discussing the labor market effects (Section 5.4.2) as well as the results for the robustness checks (Section 5.4.3), effect heterogeneity (Section 5.4.4) and possible mechanisms (Section 5.4.5).

[21] A common support statistic of our estimated models is provided in Table 5.B.3 in Appendix 5.B.

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

[23] For computations, we used STATA and the corresponding command *radiusmatch* as developed by Huber, Lechner, and Steinmayr (2015). Our specification follows mainly the default options, i.e. we employed a linear bias correction as well as a multiplier of 300 and a quantile of 90, indicating that the radius is equal to 300 percent of the maximum distance in pair matching and the 90th quantile of the distances in pair matching are used. All analyses had to be conducted in the Research Data Center of the SOEP in Berlin, since regional information in the SOEP is subject to severe data restriction policies. Unfortunately, the infrastructure at the DIW did not allow us to employ more advanced machine learning estimators.

[24] In this regard, one might argue about whether the number of bootstrap replications is sufficiently large. Again, the infrastructure at the DIW seriously limited our estimation options. However, while we were not able to compute a larger number of bootstrap replications for all the different models estimated, we compared the results of one strata (men, 2008) between bootstrapping 99 and 499 times. Overall, the results did not change much. This is in line with some recent, more general evidence suggesting that reliable inference can already be obtained with very small numbers (there: 49) of bootstrap replications (see Bodory, Camponovo, Huber, & Lechner, 2018). Therefore, we are not concerned about any serious problem in this regard.

5.4.1 Selection process

Table 5.1 provides an overview on conditional mean values as well as the marginal effects of all regional characteristics and the individual labour market indicators used in the selection models based on the main specification.[25]

The data show that several mean values of the variables differ significantly between the three samples, which underscores the relevance of controlling for them in the first stage. A closer examination of the six models reveals that, in particular, the high versus low as well as the high versus medium samples differ significantly while these differences are less evident when comparing the medium versus low samples. This is reflected in comparably lower Pseudo R^2 of the probit models for the medium versus low samples. Moreover, the observable patterns do not differ much between the male and female samples, suggesting that women and men with comparable characteristics are equally distributed in all subsamples.

With the exception of the medium versus low comparison, the data show that the larger the BIK region, the higher the level of SPE. Moreover, the higher the population density (share of unemployed in the working age population), the higher (lower) the probability of observing medium compared to high or low SPE levels. In addition, the higher the share of foreigners in the municipality, the higher the level of SPE, which might reflect the policy in line with the claim that sports can serve as a vehicle for integration. The coefficients are, however, only significant for women. Furthermore, weather conditions and location influence the level of SPE as expected. Expenditures are increasing with higher precipitation rates and municipalities in the northern states spend comparably less than North Rhine-Westphalia, whereas municipalities in the western and southern states spend comparably more. Moreover, the probability of observing low SPE levels in a municipality is lower in the eastern states compared to North Rhine-Westphalia. Finally, although individual labor market indicators suggest that living in municipalities with comparably higher SPE levels is correlated with comparably more success in the

[25] For calculating the ATEs we needed to re-estimate all probit models for each year since we have an unbalanced panel with a decreasing number of observations due to panel attrition as discussed in Section 5.3.1. While all individual characteristics as described in Section 5.3.2 are included in all models, Table 5.1 only reports results for the individual labour market indicators which are in the focus of our study. In general, we observe only few significant differences with regard to the individual characteristics (all results are reported in Table 5.B.1 in Appendix 5.B).

Table 5.1: Mean values and marginal effects of the selection models (selected variables).

	Men 20-52 years in 2001						Women 20-52 years in 2001					
	Mean in subsample			Marginal effects			Mean in subsample			Marginal effects		
Characteristics	High	Med	Low	H-L	H-M	M-L	High	Med	Low	H-L	H-M	M-L
(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)
Population size and density												
Population size $\geq$ 500k	0.44	0.38	0.32	R	R	R	0.46	0.41	0.32	R	R	R
Population size 100k-499k (L)	0.18	0.19	0.12	-0.20**	-0.53***	0.15	0.17	0.18	0.11	-0.04	-0.27***	0.05
Population size 50k-99k (M)	0.02	0.04	0.07	-0.09	-0.41	0.13	0.03	0.04	0.08	-0.002	-0.19	0.02
Population size 0-50k (S)	0.35	0.38	0.49	-0.11**	-0.42***	0.17**	0.34	0.37	0.49	-0.11*	-0.37***	0.10*
Population density	1.8	1.7	1.2	0.02	-0.14***	0.10***	1.8	1.8	1.3	0.02	-0.10***	0.06***
Population density x L	0.38	0.35	0.17	0.16***	0.22***	0.03	0.32	0.34	0.17	0.08*	0.10**	0.07**
Population density x M	0.04	0.07	0.10	-0.05	0.09	-0.06	0.04	0.07	0.11	-0.09*	-0.008	-0.03
Population density x S	0.29	0.22	0.17	0.08***	0.18***	-0.02	0.29	0.21	0.20	0.07**	0.17***	-0.008
Population characteristics												
Share of unemployed [1)]	7.1	8.9	9.2	-0.02***	0.02	-0.03***	7.2	8.9	9.4	-0.02***	0.01***	-0.03***
Share of graduates with HEEQ [1)]	24	25	23	0.003	-0.0003	0.005**	24	25	23	0.0009	-0.002	0.002
Share of foreigners	1.9	1.5	1.2	0.03	0.03	0.005	1.9	1.5	1.2	0.06**	0.02	0.04**
Weather conditions												
Average precipitation [2)]	2.1	2.0	2.0	0.33***	0.28***	0.06	2.1	2.0	2.0	0.29***	0.25***	0.02
Average sunshine duration [2)]	0.47	0.46	0.45	0.98***	0.12	0.80**	0.47	0.47	0.46	0.31	-0.51	0.34
Region												
North-Rhine Westphalia	0.22	0.20	0.21	R	R	R	0.25	0.20	0.20	R	R	R
Southern Germany	0.4	0.18	0.17	0.10**	0.20***	-0.05	0.38	0.19	0.16	0.12***	0.22***	-0.02
Northern Germany	0.04	0.14	0.20	-0.20***	-0.27***	-0.03	0.03	0.13	0.21	-0.29***	-0.31***	-0.09***
Western Germany	0.19	0.12	0.08	0.20***	0.14***	0.13**	0.19	0.12	0.09	0.16***	0.15***	0.07
Eastern Germany	0.14	0.35	0.34	0.13**	-0.17**	0.19***	0.15	0.36	0.34	0.17***	-0.16**	0.27***
Labor market indicators												
Household net income in EUR	2733	2577	2592	0.000009	0.00001	-0.00001	2640	2568	2505	0.00002*	0.000005	0.00001
Gross earnings in EUR	2731	2424	2271	0.000007	0.00002	0.00001	1165	1165	1086	-0.00002	-0.00002	0.000006
Working time per week	41	39	38	0.0008	0.001	-0.0003	23	23	23	0.0006	0.002**	-0.001
Hourly wage in EUR	14	13	12	-0.0003	-0.004	-0.0005	8	8	8	0.003	-0.001	0.003
Fulltime employed	0.88	0.84	0.81	0.02	-0.009	0.04	0.39	0.39	0.38	0.019	0.006	-0.0008
Nr of observations, Pseudo R^2	985	1,051	1,009	22	13	7	1,095	1,150	1,135	21	12	8

Note: All figures are based on the 2007 sample. For calculating the ATEs we needed to re-estimate all probit models for each year since we have an unbalanced panel with a decreasing number of observations due to panel attrition (results are available upon request). The 'high' ('med') ['low'] sample consists of individuals living in a municipality with high (medium) [low] levels of distance weighted per capita sports facility expenditures as defined in Chapter 5.2. The dependent variables in the probit models are dummies measuring one for the relatively higher expenditure level, i.e. for the H-M-model 'high', for the H-L-model 'high' and for the M-L-model 'med'. Independent variables are measured prior to the dependent variable if not stated otherwise, i.e. 1) measured between 2001 and 2006, 2) measured between 1981 and 2010. $***, **, *$ denote significance at the 1% , 5% , 10% levels. R denotes the reference category. A detailed description of all variables is provided in Table 5.B.1 in Appendix 5.B. A full length Table 5.B.2 in Appendix 5.B displays the results for all variables including the individual characteristics not displayed here. Sources: RDC of the Federal Statistical Office and Statistical Offices of the Länder (2018a; 2018b) and SOEP as well as further data sources as discussed in detail in Chapters 5.2, 5.3 and in Appendix 5.A, own calculations.

labor market, only 2 of 30 marginal effects are significant, in line with the observation that few significant differences in individual characteristics can be observed (see Table 5.B.2 in Appendix 5.B).

5.4.2 Labor market effects

We next turn to the estimated ATEs as described in Section 5.3.3. Figure 5.2 shows the corresponding effects on an annual basis of moving from low to medium, from low to high and from medium to high SPE levels for monthly (a) net household income, (b) net earnings and (c) gross earnings. To see how these effects develop and whether (and to what extent) controlling for our confounding variables was successful, we also include the corresponding estimates for the pre-exposure (2001) and exposure period (2002-2006) in each graph.

Overall, the intuition is that balancing out the different confounding influences worked the better the closer the estimated ATEs were to zero in 2001, and a negative (positive) deviation from zero indicates a potential downward (upward) bias of our estimates. Accordingly, the balancing worked slightly better for men than for women with regard to household income and vice versa with regard to earning measures.

As Figure 5.2 shows, although we do not observe any significant effects on income and earnings for moving from low to medium SPE levels (see dotted lines), we find that men (women) have an average[26] gain in net household income of around €180 (€210) when moving from medium to high (dashed lines) and €140 (€80) when moving from low to high SPE levels (solid lines). The latter effects (high versus low), however, are not significant, particularly for the female stratum.[27]

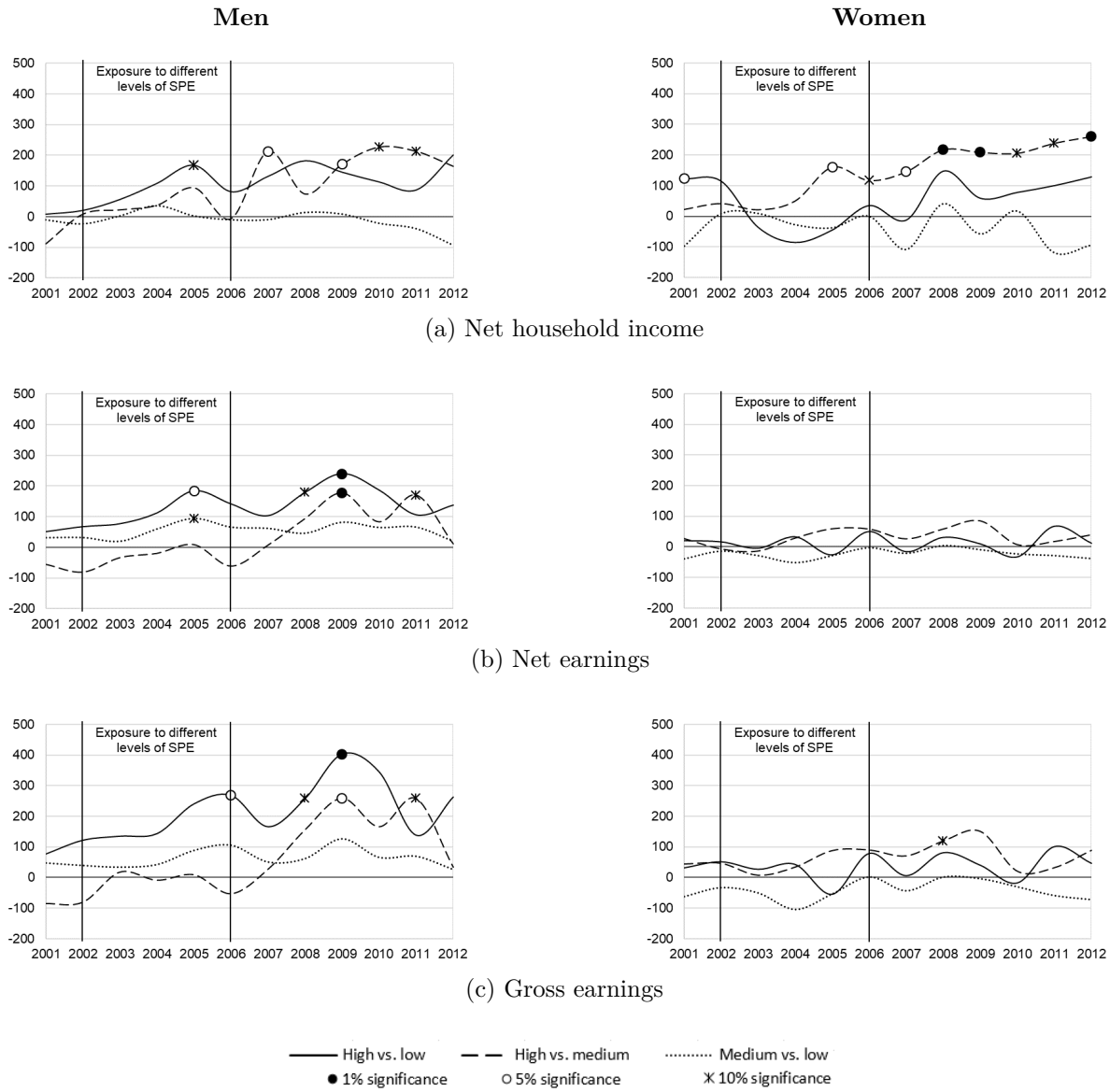
If we consider the differences in the level of balancing in 2001 between men and women, the effects of moving from medium to high SPE levels seem to converge between gender at approximately €190-€200.[28] By contrast, however, significant gender differences occur with regard to the estimated effects on earnings. Whereas men have an average gain in net (gross) earnings of around €90 (€150) when moving from medium to high SPE levels, the corresponding effects size for women is less than half and (with one exception) not significant. Moreover, if we consider again the differences in the level of balancing

[26] We use the term 'average' here and in the following discussion to refer to the six-year average in the post-exposure period (2007–2012), if not stated otherwise.

[27] As can be seen Tables 5.B.4 and 5.B.6 in Appendix 5.B, the mean and median biases are significantly larger for most covariates in the high versus low (H-L) models compared to the high versus medium (H-M) models.

[28] The 2001 ATEs for net household income show a negative deviation from zero (indicating a downward bias) for men while the corresponding deviation for women is positive (indicating an upward bias).

Figure 5.2: Effects of different levels of SPE on household income and earnings.



Note: Radius matching results for men and women. Displayed are Average Treatment Effects (ATEs) of 'high vs. low', 'high vs. medium' and 'medium vs. low' levels of distance weighted per capita sports facility expenditures (as defined in Chapter 5.2) on household income and earnings. All 2001 ATEs are calculated as mean differences between matched treated and untreated observations coming from the balancing tests of the 2002 estimations. The effects for net earnings in 2001 are approximated by multiplying the effects for gross earnings with the samples' average net-to-gross earnings ratio, i.e. 65.6% (63.5%) for men (women). Common support statistics are displayed in Table 5.B.3 in Appendix 5.B. An overview on the balancing statistics of all covariates for the estimated matching models from the post-exposure period (2007-2012) is provided in Tables 5.B.4 and 5.B.6 in Appendix B. Sources: RDC of the Federal Statistical Office and Statistical Offices of the Länder (2018a; 2018b) and SOEP as well as further data sources as discussed in detail in Chapters 5.2, 5.3 and in Appendix 5.A, own calculations.

in 2001 indicating a downward (upward) bias of results for men (women), these findings suggest that any household income effect when moving from medium to high SPE levels

is captured by earning gains for men rather than women living in the household. Our findings for moving from low to high SPE levels reinforce this interpretation: While we do not observe any significant effects for women, we observe an average gain in net (gross) earnings of approximately €160 (€260) for men. Considering that the corresponding ATEs from 2001 indicate an upward bias of our estimates for both earning measures (in contrast to the household income measure), the ATEs for household income and (net) earnings for men seem to converge by and large.

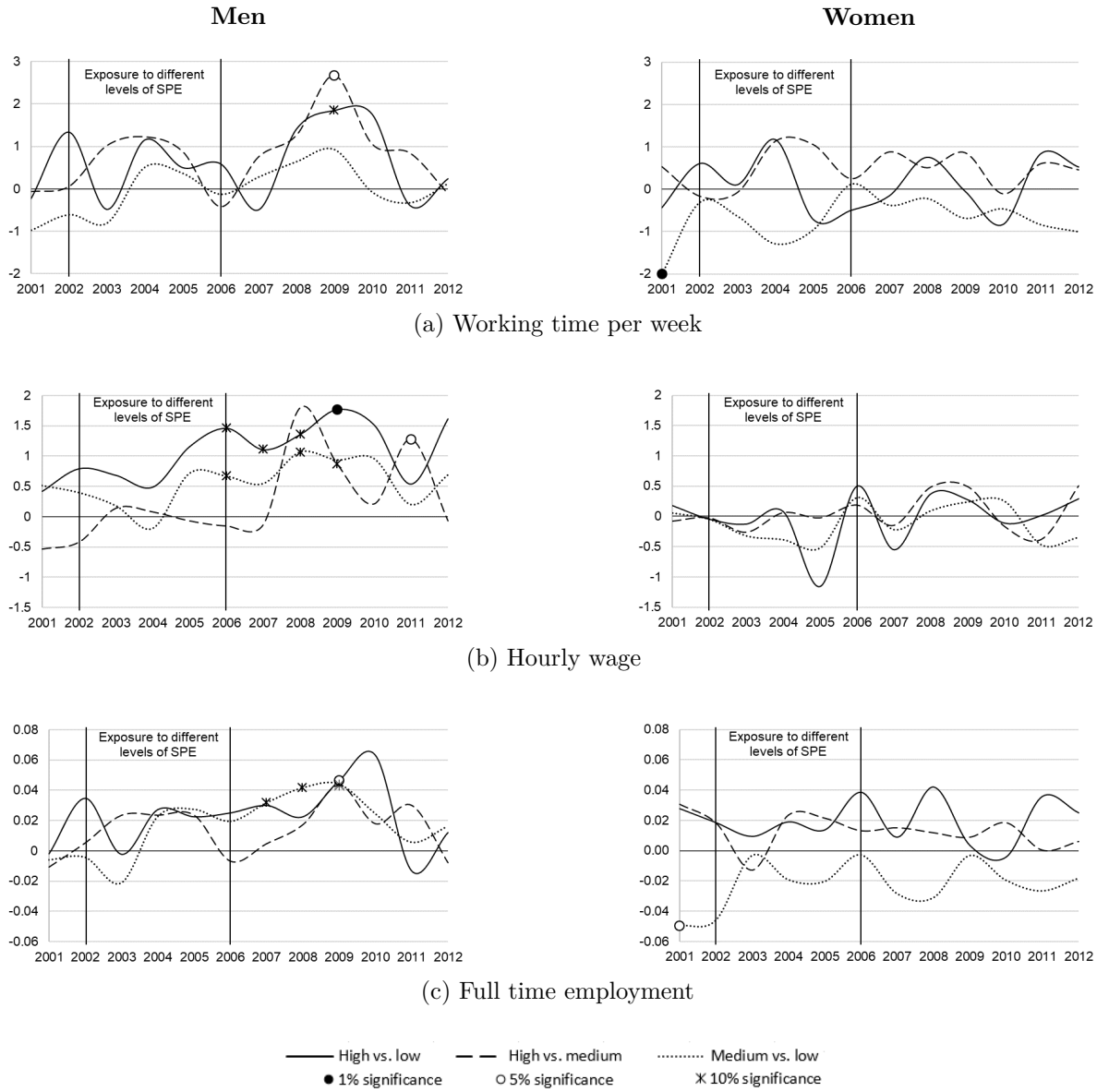
To determine whether these gender specific differences are reflected by differences in other labor market indicators as well, we also estimate the corresponding effects on working time per week, hourly wage[29] and the probability of being employed full time (see Figure 5.3).

As Figure 5.3 indicates, we do not observe any significant effects in the post-exposure period for the three measures for women, but we do find some significant effects for men. In general, however, the corresponding estimates are more volatile than the estimates for income and earnings. For example, moving from medium to high SPE levels goes along with an average increase in working time of about one hour in the post-exposure period. This effect is of comparable size for men when moving from low to high SPE levels. However, we only observe a significant estimate in 2009 for both models. Furthermore, we find a significant increase in hourly wage between (on average) €0.7 and €1.3 in the post-exposure period for all three models when moving to comparably higher SPE levels. Considering the differences in the level of balancing in 2001 between the models, these effects seem to converge at approximately €1 for the low versus high and the medium versus high comparisons. Similarly, we observe an increase in the probability of being employed full time in the post-exposure period when moving to comparably higher SPE levels. However, these effects are only significant in 2009 for the low versus high and the medium versus high comparisons.

In summary, we observe a significant (average) increase in household net income of about €190 in the post-exposure period when moving from medium to high SPE levels. These effects are, however, fully captured by earning gains for men rather than women living in the household. Moreover, while our estimates for other labor market indicators

[29] Because hourly wage is not available in the SOEP, it is approximated by dividing monthly gross earnings by weekly hours (x 4.3). For non-workers, hourly wage is zero.

Figure 5.3: Effects of different levels of SPE on working time, wages and employment status.



Note: Radius matching results for men and women. Displayed are Average Treatment Effects (ATEs) of 'high vs. low', 'high vs. medium' and 'medium vs. low' levels of distance weighted per capita sports facility expenditures (as defined in Chapter 5.2) on hourly wages, employment status and working time per week in hours. All 2001 ATEs are calculated as mean differences between matched treated and untreated observations coming from the balancing tests of the 2002 estimations. Common support statistics are displayed in Table 5.B.3 in Appendix 5.B. An overview on the balancing statistics of all covariates for the estimated matching models from the post-exposure period (2007-2012) is provided in Tables 5.B.4 and 5.B.6 in Appendix 5.B. Sources: RDC of the Federal Statistical Office and Statistical Offices of the Länder (2018a; 2018b) and SOEP as well as further data sources as discussed in detail in Chapters 5.2, 5.3 and in Appendix 5.A, own calculations.

are generally quite volatile, the gender differences we observe for earnings are broadly confirmed. While we do not find any significant effects for women in the post-exposure period, our results suggest that (on average) men work about one hour more and benefit of a €1 increase in wage when moving to comparably higher SPE levels.

5.4.3 Robustness checks

To establish the robustness of these findings, we re-estimate our models with different samples and different specifications of both measures and models. First, we test whether our results in the exposure periods are driven by individuals who were still in the education system between 2001 and 2006, because (for example) students are not randomly distributed across municipalities.[30] While the available regional and individual confounders are expected to account for this issue, empirical testing seems advisable since the estimated ATEs in the (pre-)exposure period(s) already provide important information about the balancing and about how effects develop over time. Red lines in Figures 5.C.1 and 5.C.2 in Appendix 5.C show the estimates for all samples excluding individuals who were still in the education system in 2001-2006 (e.g., trainees, students, civilian servants, conscripts; approximately 10-11 % of the sample). Although some minor deviations in effect sizes emerged and balancing out the different confounding influences for the low versus high comparison worked better when individuals still pursuing education are included, the results do not indicate that our main findings are affected.

An additional robustness check considers the identification strategy employed. A natural candidate to be included in our selection models is SPE measured in the pre-exposure year 2001. Controlling for lagged SPE is certainly a comparably stricter condition that might add more credibility to our identification strategy. However, we avoided this approach as a main specification because it requires a sufficiently large variation between the 2001 and the (average of the) 2002-2006 SPE-levels. Consequently, this approach emphasizes comparisons between treatment and controls close to the thresholds (i.e., between medium versus low and high versus medium SPE levels). Moreover, it reduces the probability of finding adequate matching partners for the extreme case, that is, the

[30]Note that we only excluded individuals who were still in the education system in 2007 or later in the main specification in order to keep as many observations as possible.

comparison between high and low SPE levels.[31] Figure 5.C.3 in Appendix 5.C displays the estimates for the male panel in the post-exposure period including lagged SPE.[32] As expected, we only observe a few significant estimates as a result of a considerable loss in precision. Moreover, the average gains in net household income [approximately €140 (€120)], gross earnings [approximately €150 (€120)] and net earnings [approximately €90 (€80)] when moving from medium (low) to high SPE levels decrease, particularly for the latter comparison (i.e., low versus high SPE levels). With regard to the other labor market indicators, a similar picture emerges.

We also test the results when using our original 'pure' SPE measure instead of the radius measure as defined in Chapter 5.2. As could be expected, the results (displayed in Figure 5.C.4) do differ (in some parts, considerably) from the results we derived when using our radius measure. We still observe some positive gains in net household income (average effect: €170) and net earnings (average effect: €130) when moving from low to high SPE levels that are comparable in size to our estimates using the distance weighted radius measure. However, none of the income and earning ATEs for the other two comparisons (high versus medium or medium versus low) are significant. Moreover, only the effect on full-time employment is significant (and of similar size compared to our radius measure estimates) for the medium versus low comparison.

Finally, we test whether panel attrition may be an issue by checking whether being in either of the SPE groups leads to selective attrition, which could potentially invalidate our estimates. To do so, we make use of the full dataset after deleting all observations with missing values for our confounding variables and estimate the effects of SPE level on a binary outcome variable, which equals 1 if there is missing information in any outcome variable of interest of the respective year for a particular individual. Note that such a test goes beyond a pure panel attrition test and is somewhat stricter since individuals also drop out of our analysis in years where we do not observe any outcome (although they may still remain in the panel). While only 1 out of 10 estimates is significant at the 10 percent level for women when moving from medium to high SPE levels, 4 out of 10

[31] This is confirmed by the common support statistics indicating comparably more observations off support (see Table 5.B.3 in Appendix 5.B) as well as the balancing statistics indicating comparably higher mean and medium biases for most covariates in particular for the H-L models (see Table 5.B.4 in Appendix 5.B).

[32] Due to computational constraints imposed by the infrastructure at the DIW, we were not able to estimate all robustness checks for the pre-exposure and exposure periods.

estimates for men and women are significant when moving from low to high SPE levels.[33]

In summary, while excluding individuals who were still in the education system between 2001 and 2006 changes hardly any findings for the exposure period, both controlling for lagged SPE in the selection process as well as using the pure (instead of our radius) SPE measure does change our findings for the post-exposure period. Importantly, however, even for these specifications we observe some positive labor market effects when moving to comparably higher levels of SPE. Moreover, while controlling for lagged SPE in the selection process leads to an expected loss in precision, using the pure SPE measure does not appropriately capture the situation for individuals who may arguably benefit from sports infrastructure in their neighboring municipalities. Therefore, these results remain as a robustness check in our analysis. Finally, while we cannot fully rule out any panel attrition bias in general, our test results suggest that at least the main findings for the medium versus high comparison are hardly affected.

5.4.4 Effect heterogeneity

From a policy perspective, it is highly relevant to understand whether certain societal groups benefit more or less than others. As mentioned previously, significant differences occur in terms of gender. While men and women exposed to comparably higher levels of SPE benefit from an increase in household income, additional analysis reveals that these gains in income are captured by earning gains for men rather than women living in the household. To determine whether this result is plausible, we need to more closely investigate men and women's differing sports participation habits. To do so, we make use of the SOEP-Innovation Sample (SOEP-IS) which covers more than 3,000 households. It was established in 2012 (Richter & Schupp, 2015) and includes a module on physical activity in the years 2013, 2015 and 2017.[34] According to these data, around 38% of women (35% of men) in our age cohort (20-52 years) did not engage in any sports. Within our observation window (2013-2017), the share of inactive men remained rather stable while the corresponding share of inactive women increased. As a consequence, the gender gap

[33]Results are available upon request.

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

in terms of inactivity increased to about five percentage points within four years. We also observe significant gender differences in the settings where sport is practiced. In this regard, 16 (12) percent of women (men) engage in their first sport with a commercial sports provider. This is in line with the finding that 56 percent of members in fitness centres in Germany are female (Deloitte, 2018). Overall, these figures suggest, that women are generally less likely to engage in sports than men. Moreover, active women have a higher probability of engaging in sports outside any of the publicly funded sports facilities. Therefore, it appears plausible to assume that women have a lower probability to benefit from higher levels of SPE than men.

In order to test whether any of the observed labor market effects depend on age, we split the male sample into two groups (men aged 20-36 and men aged 37-52 in 2001). The results for the post-exposure period are displayed in Figure 5.D.1 in Appendix 5.D. While we do not find any effects for the younger cohort, we do find several significant effects for the older cohort. With regard to net household income and compared with estimates for the full male sample, these effects are larger (€220) for the high versus low SPE groups and of similar size (€170) for the high versus medium SPE groups. With regard to the other labor market indicators, we find an average positive effect of an increase in working time (1.4 hours), hourly wage (€1) and on the probability of being full time employed for the high versus low comparison. The corresponding effects for the high versus medium comparison are, however, not significant (with the exception of working time in 2009). Again, we use the SOEP-IS data to shed light on whether these age specific differences for men seem plausible. According to the data, around 24 percent of the younger cohort (men aged 20-36 years) did not engage in any sports. This share is considerably higher (44 percent) for the older cohort (men aged 37-52). Although the share of inactive men increased during our observation window (2013-2017) by 2 percentage points in the younger cohort while it decreased by 3 percentage points in the older cohort, the gap is still considerable. At the same time, however, younger cohort members have on average a 5-percentage-point higher probability than older cohort members (15% vs. 10%) of engaging in their first sport with a commercial sports provider. Therefore, although younger cohort members have a higher probability of practicing sports in general, they are more likely to do so outside any of the publicly funded sports facilities. We thus

conclude that it is not unlikely that younger men have a lower probability of benefiting from higher levels of SPE than older men.

In summary, we observe some effect heterogeneity by gender and age. While both men and women exposed to comparably higher levels of SPE benefit from an increase in household income, these gains are captured by earning gains for men rather than women living in the household. Further analysis reveals that these effects unfold more for older (aged 36-52 years) than younger (aged 20-36 years) men. While both findings are supported by some descriptive figures on gender and age differences with regard to sports participation habits coming from the SOEP-IS, it is important to note that the disaggregated analysis by age strata suffers from an efficiency loss due to comparably smaller sample sizes.[35]

5.4.5 Exploring mechanisms

The underlying assumption of our analysis is that differences in per capita expenditures for sports facilities between municipalities are associated with differences in the quantity and/or quality of sports facilities. Although we are unable to test this assumption at the local level given a lack of data, we do find some supportive evidence at the regional level by correlating the average level of SPE with different indicators for sports facility provision in the 16 states ('Länder').[36] Because greater access to sports facilities increases sports participation (Huang & Humphreys, 2012) and higher levels of exercise in turn improve productivity which is rewarded in the labor market (Lechner, 2009; Lechner & Sari, 2015), we argue that the observed labor market effects may unfold from the SPE induced differences in sports participation behavior.

To test this assumption directly, we used the continuous nature of our SPE measure as well as the available sports participation measure in the SOEP and estimated dose-

[35] We also tested whether any labor market effects might depend on whether children are living in the household by re-estimating our models for men living with / without children in the household separately. Our assumption was that comparably higher levels of SPE might positively affect the participation patterns of children which in turn might enable parents to increase their productivity at work. However, we do not find this relationship in the data.

[36] We find some positive correlations between the average level of SPE and the total number of sports facilities as well as the area (in square meters) covered by outdoor and indoor pools. The correlation between the average level of SPE and the provision with area covered by indoor pools (in square meters per 1,000 inhabitants) is significant at the 10 percent level (see Table 5.E.1 in Appendix 5.E).

response functions with various specifications.[37] While we could not find this expected relation in our data, it is important to note that the available measure in the SOEP only considers the *frequency* of sports participation and does not measure any differences in *intensity* and/or *duration* of sports participation. A certain composition of all three characteristics describing sports participation is, according to the World Health Organization (World Health Organization, 2010), however, required for achieving positive health effects.[38] Therefore, we performed another more indirect test for differences in sports participation behavior by testing for differences in well-being and health. To do so, we estimated and compared the effects for men between our main specification and the main specification including satisfaction with health, satisfaction with life, the number of doctor visits per year as well as the physical component (PCS) and the mental component summary scales (MCS) from the corresponding year as additional covariates in the selection process. The idea is that any labor market effects caused by an increase in health enhancing sports participation should disappear when controlling for differences in health status.[39] Figure 5.E.1 in Appendix 5.E displays the effects of different levels of SPE on household income, earnings, working time, wages and full time employment for men during the post-exposure period excluding (main specification, black lines) and including (red lines) our health measures from the corresponding year as additional covariates in the selection process. Visual inspection of these graphs suggests, that some effects indeed decreased when controlling for the various health measures. This reduction is more striking for even years (2008, 2010 and partly 2012), for which we could control for the most sophisticated health measures available, i.e. the PCS and MCS.

Table 5.E.2 in Appendix 5.E tries to quantify the overall effect of controlling for the health measures. The table displays the ATEs for both specifications (ATE and ATE_{Health}) and the average change (in percent) between ATE and ATE_{Health} calculated for all years in the period 2007-2012 ($Change^1$) as well as only for years with significant estimates ($Change^2$).[40] After controlling for the different health measures, we found a

[37] Results are documented in a separate paper (see Steckenleiter et al., 2019).

[38] Supported by various studies offering empirical evidence for this, the WHO guidelines suggest that 150 min (75 min) per week of moderate (vigorous) regular physical activity is required in order to achieve any positive health effects.

[39] Data for constructing the PCS and MCS measures is only gathered in even years since 2002 and is therefore not available for 2007, 2009 and 2011.

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

reduced effect for net household income and earnings of approximately 15 percent as well as a reduced effect for wage (being employed full time) of approximately 8-10 (6) percent.[41] Importantly, when controlling for PCS and MCS is possible, these values increase to around 40 percent for income and earnings and approximately 20 percent for wage and being employed full time.

Overall, these findings suggest that the estimated positive effects on labor market outcomes might indeed be driven by health enhancing sports participation, though other channels beyond health remain relevant. In this regard, Lechner (2009) notes that such channels might relate to, for example, social networking. Indeed, using SOEP data, Schüttoff, Pawlowski, Downward, and Lechner (2018) show that regular sports participation positively affects adolescents' social capital through volunteering, helping friends, and civic involvement. Unfortunately, however, information on these measures is only available for uneven years and thus not available for 2008, 2010 and 2012, which makes any joint effect analysis together with all health measures impossible.

5.5 Conclusion

By merging administrative data of all municipalities in Germany with individual data coming from the SOEP, we explore whether local public expenditures on sports facilities influence individual labor market outcomes. We approach the issue that public expenditures are policy variables and individuals may deliberately make location choices based on the availability of sports infrastructure by exploiting the panel structure of our data and controlling for variables expected to jointly influence public expenditures and labor market outcomes. To this end, we use several regional characteristics to control for any confounding influences at the *aggregate* (municipality) level and several individual characteristics to control for any confounding influences at the *disaggregated* (individual) level.

Overall, our matching analysis reveals sizable effects of public expenditures on individual labor market outcomes. For both men and women, we do observe a significant (average) increase in household monthly net income of about €190 (approximately 7% of

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

average household income[42]) when moving from medium (€20-€31) to high (€31-€85) levels of SPE. This translates into an average increase in annual spending of about €2.5 million for a medium sized city with 75,000 inhabitants. Interestingly, however, these effects are captured by earning gains for men rather than women living in the household. Additional analysis confirms and supports these apparent gender differences.

Whereas we do not observe any significant effects for women, we find some evidence suggesting that (on average) men work one hour more and benefit by a €1 increase in hourly wage as well as an increase in the probability of being employed full time when moving to comparably higher SPE levels. Moreover, descriptive figures of the SOEP-IS show that women in the age cohort under consideration are less likely than men to engage in sports in *general* and in any of the publicly funded sports facilities *in particular*. We also find that older (aged 37-52 years at the beginning and 48-63 years at the end of our observation window) rather than younger (aged 20-36 years at the beginning and 31-47 years at the end of our observation window) men benefit from higher SPE levels. This result is also supported by some descriptive figures coming from the SOEP-IS data suggesting that the younger cohort has a comparably higher probability of engaging in sports outside any of the publicly funded sports facilities. Finally, although possibilities to test whether these effects unfold due to (SPE induced) differences in sports participation patterns are limited, we find some evidence that improved well-being and health are possible mechanisms.

With this study, we intend to contribute to the rich literature analyzing the effects associated with public expenditures, which to date has largely neglected investigating types of public spending that might indirectly improve health, educational and labor market outcomes. Unfortunately, however, while health, education and labor markets are clearly defined settings, sport and sports participation are not, which results in some imprecision in our analysis. For example, in contrast to Jackson et al. (2016) and others who examine the effects of local school spending, we have difficulties in specifying the exposure period required for the effects we found to unfold. Although our findings suggest that controlling for exposure to different levels of SPE over five years is sufficient to reveal positive effects, specifying more precisely the conditions under which these effects unfold

[42] Average household net income is calculated as frequency weighted average of the six income figures for 2007 as displayed in Table 5.1.

requires more attention in future studies.

Moreover, even the more established literature from settings such as health, education and labor markets remains rather silent about opportunity cost arguments to date – probably because health, education and labor markets are topics that are in the policy focus of all local, regional and national governments around the world. This, however, is not the case for sports (at least not to a similar extent), for which the budget is frequently in competition for scarce public money available for other 'cultural' areas such as museums or theaters. In this regard, it would be relevant to test whether and to what extent public spending on other areas beyond health, education and labor markets might be associated with any effects on labor market outcomes as found for sports in this study. However, whether this will be possible in the future depends not only on data quality but also on data protection rules and regulations, which have also considerably limited the scope for implementing more comprehensive econometric techniques in our study.

5.A Data cleaning, transforming, and enriching process

The Scientific Use Files (SUFs)[43] for the years 2001 to 2006 consist of information on all four different types of local authorities: urban municipalities/cities *('kreisfreie Städte')* and rural districts *('Landkreise')* which comprise more than 12,000 municipalities *('Gemeinden')* the majority of which are organized in municipality associations *('Gemeindeverbände')*. In our analyses, we focus on the lowest administrative units only, that is, urban municipalities/cities *('kreisfreie Städte')* and district municipalities *('kreisangehörige Gemeinden')*, since budgets of the municipality associations are commonly made up by the affiliated municipalities' budgets and sports related expenditures by rural districts are negligibly small.[44] The SUF data are available as long format only. This means each revenue and expenditure group for each MC/subcategory for each local authority makes a single observation, summing up to more than 5 million revenue/expenditure group-MC/subcategory-municipality-observations per year. To work with it, we transformed the data into a wide format with each type of local authority as single observation.

Next, the municipality key codes (AGS) were generated for the years 2001 to 2012 and pooled with the RDB as they build the unique identifier for further pooling our data with local level information coming from other sources. The original key code available in the SUF *('Berichtsstellenschlüssel')* consists of the codes of the administrative district *('Regierungsbezirk')* (one digit), the rural district (two digits), the last two digits of the municipalities association (two digits), and the code for the municipality (three digits). The AGS, however, consists of eight digits which comprises information on the federal state (two digits), the administrative district (one digit), the rural district (two digits), and the municipality (three digits) only.

During these transformations, the data were cleaned. For the municipalities in Saxony–Anhalt (ST) the codes in the SUF were outdated and had to be updated with the help

[43] Official reference to the data: RDC of the Federal Statistical Office and Statistical Offices of the Länder (2018b).

[44] Some municipality associations recorded sports related incomes and expenditures. Since, however, the budgets of municipality associations are made up by the affiliated municipalities and we operate with gross values of municipality accounts (see explanations further below), all sports related incomes and expenditures of municipality associations are already included in the municipality accounts which are used in our analysis.

of the municipal directories. Also, some urban municipalities and district municipalities merged over time and have the same AGS in later years. Overall, the number of municipalities decreased from N=12,930 in 2002 to N=12,291 in 2006 (see Table 5.A.1).

Table 5.A.1: Development of the number of municipalities between 2002 and 2006 in the RDB

Federal state	**2002**	**2003**	**2004**	**2005**	**2006**
Baden-Wuerttemberg (BW)	1,111	1,111	1,111	1,111	1,110
Bavaria (BY)	2,056	2,056	2,056	2,056	2,056
Berlin (BE)	1	1	1	1	1
Brandenburg (BB)	886	436	421	420	420
Bremen (HB)	1	1	1	1	1
Hamburg (HH)	1	1	1	1	1
Hesse (HE)	426	426	426	426	426
Lower Saxony (NI)	1,022	1,022	1,022	1,022	1,021
Mecklenburg-Western Pomerania (MV)	979	964	873	851	849
North Rhine-Westphalia (NW)	396	396	396	396	396
Rhineland-Palatinate (RP)	2,306	2,305	2,306	2,306	2,306
Saarland (SL)	52	52	52	52	52
Saxony (SN)	535	525	519	514	510
Saxony-Anhalt (ST)	1,027	1,027	1,027	1,027	1,027
Schleswig-Holstein (SH)	1,124	1,123	1,123	1,123	1,123
Thuringia (TH)	1,007	1,006	998	998	992
Total	**12,930**	**12,452**	**12,333**	**12,305**	**12,291**

Source: RDC of the Federal Statistical Office and Statistical Offices of the Länder (2018b), own calculations.

To generate a balanced panel, information on merging municipalities was manually aggregated also for earlier waves. A comparison of the RDB with the municipal directory of the corresponding years indicated some differences with regard to the overall numbers of local authorities in Germany for some years. Such differences occurred for two reasons: 'premature mergers' and 'belated mergers'. 'Premature merger' of municipalities in the RDB occur since some municipalities that are listed as single municipalities in the municipal directory were already aggregated in the RDB (examples include *Bitterfeld-Wolfen, Dessau-Roßlau, Elsteraue, Bördeland*). Since disaggregated sport related expenditures for this premature merger municipalities were not available, we decided to use the already merged observations. 'Belated merger' of municipalities in the RDB occur since some municipalities that are listed as single municipalities in the RDB were already aggregated in

the municipal directory (examples include several municipalities in Brandenburg in 2001 such as *Brodowin* and *Serwest* which were already merged with Chorin in the municipal directory). This issue was fixed by a subsequent aggregation of the 'to-be-merged' municipalities. To facilitate further analyses, we harmonized the data base so that for every municipality only one single observation is left. After reformatting, further data checks were applied and any occurring inconsistencies were removed.

Overall, eight variables are included in the SUFs: EF1: Municipality key code *('Amtlicher Gemeindeschlüssel – AGS')*, EF1U1: Region *('Bundesland')*, EF2: Year *('Erhebungsjahr')*, EF3: Figures in Euros or number of inhabitants *('numerische Angaben')*, EF4: Classification number *('Gliederungsnummer Bund')*, EF5: Grouping number *('Gruppierungsnummer Bund')*, EF6: Municipality size range *('Gemeindegrößenklasse')*, EF7: Kind of local authority *('Gebietskörperschaft')* and EF8: Authority name *('Gemeindebezeichnung')*. The classification (EF4) and grouping (EF5) numbers are of particular interest in our project since they unambiguously identify the various kinds of public incomes and expenditures at the local level (including the money from regional authorities channeled through the local authorities). Our target variables are total expenditures (EF5: *'999 Ausgaben des Verwaltungs- und Vermögenshaushalts')* for sports facilities (EF4: *'56 eigene Sportstätten')* and swimming pools (EF4: *'57 Badeanstalten'*) including leisure-time sports facilities, sports airfields, sports stadiums, sports fields, tennis courts, toboggan and bobsled runs, sports schools, gyms, ski jumps and Olympic sports facilities as well as indoor and outdoor pools (Federal Statistical Office, 2016).[45]

Finally, due to the increasing heterogeneity in accounting standards between municipalities since 2007 we focus our analysis on the financial years before.[46] More precisely, we took the 5-year average of municipality expenditures on sports facilities (including swimming pools) between 2002 and 2006 converted to 2004 Euros using the Consumer Price Index from the Federal Statistical Office (2017) as basis for our expenditure measure.

Over time, an increasing number of municipalities have transformed sports facilities

[45] We did not consider the general category *'funding of sports'* (EF4: '55 Förderung des Sports') since it is very broad and imprecise including general funding of sports clubs and sports association and particular funding measures to support high-performance sport as well as leisure sport (e.g., by subsidizing trainers and coaching programs) (Federal Statistical Office, 2016).

[46] Some local/regional authorities have changed from fiscal accounting ('Kameralistik') to double-entry bookkeeping ('Doppik') making a comparison of municipalities with different accounting standards impossible.

and/or swimming pools into *owner-operated community enterprises* (OCPs, *'Eigenbetriebe'*), which incomes and expenditures are not displayed in the official accounting data of the SUFs. In order to harmonize the data base, we added information on all sports related OCPs coming from the annual balance sheets of public funds, institutions and enterprises in Germany accessed via PC workplaces in the Research Data Center of the Regional Statistical Office in Stuttgart/Baden-Wuerttemberg (RDC of the Federal Statistical Office and Statistical Offices of the Länder, 2018a). To avoid any double-accounting we reduced all OCP figures by the amount of subsidies received from municipalities since these subsidies are already included in the accounting figures of the SUFs. Due to strict non-disclosure conditions, we were only allowed to keep the accounting information of OCPs for municipalities running at least three sports related OCPs. In order to reach this threshold, we melted the original categories sports facilities (EF4: *'56 eigene Sportstätten'*) and swimming pools (EF4: *'57 Badeanstalten'*) to a single category and added 'artificial' accounting data of OCPs for those municipalities running less than three sports related OCPs. In order to minimize the measurement error for these municipalities, we randomly assigned accounting data of small OCPs (identified amongst municipalities with less than 10,000 inhabitants) and adjusted the figures of contaminated municipalities in a second step by subtracting the population weighted accounting data.[47]

The accounting data available does not allow to distinguish between high-performance and leisure sports facilities. Since the majority of sports facilities are commonly used jointly by high-performance and leisure sports participants/clubs, we do not consider this as a major problem. An exception, however, are professional soccer stadiums. While we are unable to disentangle soccer stadium related expenses in our data, we make use of the fact, that professional soccer clubs (as long as the clubs do not own the stadium) have to pay a rent to the municipalities/affiliated OCPs for staging games. In other words, comparably larger sports facility expenditures by communities/OCPs owing professional soccer stadiums are usually accompanied by comparably larger sports facility related incomes. To consider this potential bias, we focus on net expenditures only.[48] While this approach purges a potential bias due to soccer stadium related maintenance expenditures,

[47]Note, any observations living in municipalities used to adjust the figures of contaminated municipalities are deleted from our analysis.

[48]This is done by subtracting sports facility related income (income grouping number *'399 Einnahmen des Verwaltungs- und Vermögenshaushalts'* in the purchased SUF) from our expenditure measure.

it does not consider the larger investments done for the construction/renovation of soccer stadiums. This might be a concern particularly since the construction/renovation of several soccer stadiums for the FIFA World Cup 2006 falls into our observation window. Therefore, we decided to reduce the sports related public expenditures in the twelve cities hosting any of the FIFA World Cup 2006 games by the amounts spent for the construction/renovation of stadiums for this event. This information was gathered from official accounting books of the cities. As expected most of the FIFA World Cup 2006 related expenses fall into our category of interest, i.e. sports facilities (see in Table 5.A.2).

For reasons of comparability across differently sized municipalities, all figures were converted into per capita figures. Moreover, individuals may not only benefit from the expenditures (and the resulting services and facilities) of their own municipalities, but also of neighboring ones. Therefore, we developed a distance weighted expenditure measure as further explained in Chapter 5.2. To decide on the radius size r, we made use of a representative survey of citizens living in a metropolitan area in southern Germany, i.e. Stuttgart, based on a Computer Assisted Telephone Interviewing. In this survey, respondents were asked about their maximum willingness to spend travel time in order reach a facility for practising their sports. We took the average stated value of 26 minutes (Pawlowski et al., 2009) and used information from the German mobility report (infas, DLR, 2010) to convert these figures into three different distance measures.[49]

To work with these regional data, further information is merged to the information derived from the SUF: Detailed municipality characteristics such as (un-)employment-rates, proportion of recreational/open/natural spaces, migration rates and many others were added from a data base called INKAR (Indicators and maps on spatial and urban development *'Indikatoren und Karten zur Raum- und Stadtentwicklung'*) coming from the Federal Institute for Research on Building, Urban Affairs and Spatial Development *('Bundesinstitut für Bau-, Stadt- und Raumforschung – BBSR')*. We used data of the 2013 edition (BBSR, 2013). The municipality key code (*'AGS'*) of the year 2011 was used for merging the two data sets. Since data for municipalities organized in municipality associations (*'Gemeindeverbände'*) is only available at the aggregate (municipality associational) level, we assigned each municipality the (same) value from the corresponding

[49] According to this report, the average speed in core cities (densely populated suburban districts) [rural areas] is 26 (31) [33] kilometers per hour allowing to travel 11.3 (13.4) [14.3] kilometers within 26 minutes.

Table 5.A.2: FIFA World Cup 2006 related expenditures of the 12 cities hosting games between 2001 and 2006 in EUR million

City hosting games	**Subcategory general funding of sport ('*55 Förderung des Sports*')**	**Subcategory sports facilities ('*56 eigene Sportstätten*')**	***Owner-operated community enterprises***	∑	**Notes**
Berlin	7.060	33.313	0	40.373	1)
Dortmund	0	0	0	0	
Frankfurt a. M.	0	17.248	0	17.248	2)
Gelsenkirchen	0	0	0	0	3)
Hamburg	0	3.923	0	3.923	4)
Hannover	.137	30.458	0	30.595	5)
Kaiserslautern	54.621	0	0	54.621	6)
Cologne	4.013	7.849	0	11.862	7)
Leipzig	1.700	77.593	0	79.293	8)
Munich	.170	0	0	.170	9)
Nuremberg	.031	.562	2.574	3.167	10)
Stuttgart	0	60.095	0	60.095	11)
∑	67.731	231.042	0	301.348	

Notes: Real values as reported in the official accounting books of the cities in Germany; Figures were converted to 2004 Euros using the Consumer Price Index from the Federal Statistical Office (2017) in Germany after inclusion in the RDB.
1) The total amount spent by local authorities for the construction / renovation of the stadium in Berlin is zero (German Federal Government, 2006). However, reported figures in the official accounting books are not directly related to the stadium but include compensation payments for the local first division football club (Hertha BSC Berlin) to develop training facilities during the construction period as well as other expenditures related to the FIFA World Cup 2006.
2) The total amount spent by local authorities for the construction / renovation of the stadium in Frankfurt is EUR 64 million (German Federal Government, 2006). However, the budget of the operating company of the stadium is recorded in a 'non-sports related' category and therefore not considered in the regional data base of this research project. Therefore, only direct subsidies (EUR 17.248 million) are considered for adjusting the regional data base of this project.
3) The stadium was constructed before the period under consideration, i.e. 2001–2006.
4) The stadium was constructed before the period under consideration, i.e. 2001–2006. Reported FIFA WC 2006 related expenditures in the official accounting books are not related to the stadium.
5) The total amount spent by local authorities for the construction / renovation of the stadium in Hannover is EUR 24 million (German Federal Government, 2006). However, between 2001 and 2002 around EUR 7 million of stadium related expenditures were reported in addition in the official accounting books of the city.
6) The total amount spent by local authorities for the construction / renovation of the stadium in Kaiserslautern is EUR 8 million (German Federal Government, 2006). However, reported figures in the official accounting books (and thus also in the regional data base of this research project) include subsidies from regional authorities, i.e. Rhineland-Palatinate.
7) The total amount spent by local authorities for the construction / renovation of the stadium in Cologne is EUR 26 million (German Federal Government, 2006). However, the budget of the operating company of the stadium is recorded in a 'non-sports related' category and therefore not considered in the regional data base of this research project. Therefore, only other FIFA World Cup 2006 related expenditures (EUR 11.862 million) are considered for adjusting the regional data base of this research project.
8) The total amount spent by local authorities for the construction / renovation of the stadium in Leipzig is EUR 12 million (German Federal Government, 2006). However, reported figures in the official accounting books (and thus also in the regional data base of this research project) include subsidies from both regional authorities, i.e. Saxony, and the German federal government.
9) The total amount spent by local authorities for the construction / renovation of the stadium in Munich is zero (German Federal Government, 2006). However, reported figures in the official accounting books are not related to the stadium but include other expenditures related to the FIFA World Cup 2006.
10) The total amount spent by local authorities for the construction / renovation of the stadium in Nuremberg is EUR 28 million (German Federal Government, 2006). However, the budget of the operating company of the stadium is recorded in a 'non-sports related' category and therefore not considered in the regional data base of this research project. Therefore, only other FIFA World Cup 2006 related expenditures (EUR 3.167 million) are considered for adjusting the regional data base of this research project.
11) The total amount spent by local authorities for the construction / renovation of the stadium in Stuttgart is EUR 36 million (German Federal Government, 2006). However, reported figures in the official accounting books (and thus also in the regional data base of this research project) include subsidies from regional authorities, i.e. Baden-Württemberg.

Figure 5.A.1: Distribution of geometric centers (centroids) for each municipality in Germany



Notes: Left: Distribution of centroids for each municipality in Germany (green points, with the largest municipalities with regard to square kilometers being colored in turquoise). Right: Example of included (green) and excluded (red) municipalities in our distance weighted expenditure measure. Own calculations.

municipality association. Moreover, population figures as well as information on the district area in square kilometers were taken from the municipality directories (2001–2006) (Federal Statistical Office (Destatis), 2002, 2003, 2004, 2005, 2006, 2007). For Bremen and other municipalities where this information was not automatically generated in the pooling process it was manually added. Finally, meteorological data for average precipitation in millimeter, temperature in degree Celsius as well as sunshine hours were taken from the Climate Data Center of the German Weather Service (*'Deutscher Wetterdienst'* – DWD). We use data for the period 1981–2010 (DWD (Deutscher Wetterdienst), 2016). After deleting extreme locations (such as weather stations on mountains like Zugspitze or Brocken) municipalities were pooled with this information coming from the closest weather station based on geographic coordinates by using the Stata-command *geonear*.

5.B Variable descriptions and further estimation results

Table 5.B.1: Variable description

Variables	Description	Type
SPE	Sports facility expenditures between 0 and percentile 99	Metric
Population size and density		
Population size $\geq$ 500k (R)	Population in 2001: 500,000 and more inhabitants	Dummy
Population size 100k-499k (L)	Population in 2001: 100,000 – 499,999 inhabitants	Dummy
Population size 50k-99k (M)	Population in 2001: 50,000 – 99,999 inhabitants	Dummy
Population size 0-50k (S)	Population in 2001: 0 – 49,999 inhabitants	Dummy
Population density	Population per km^2, divided by 100	Metric
Population density x L	Population in 2001 x Population per square km	Metric
Population density x M	Population in 2001 x Population per square km	Metric
Population density x S	Population in 2001 x Population per square km	Metric
Population characteristics		
Share of unemployed	Share of unemployed in the working age population (mean of 2001–2006)	Metric
Share of graduates with HEEQ	Share of graduates with HEEQ ("Hochschulreife") (mean of 2001–2006)	Metric
Share of foreigners	LN of share of foreigners (Census 2011)	Metric
Weather conditions		
Average precipitation	LN of precipitation in mm; average sum per year 1981–2010, divided by 100	Metric
Average sunshine duration	LN of sunshine duration in hours; average per year 1981–2010, divided by 1000	Metric
Region		
North-Rhine Westphalia (R)	North-Rhine Westphalia (yes = 1)	Dummy
Southern Germany	Baden-Wuerttemberg and Bavaria (yes = 1)	Dummy
Northern Germany	Bremen, Hamburg, Lower-Saxony, Schleswig-Holstein (yes = 1)	Dummy
Western Germany	Hessen, Rhineland-Palatine, Saarland (yes = 1)	Dummy
Eastern Germany	Berlin, Brandenburg, Meckl.-West Pom., Saxony, Saxony-A., Thuringia (yes = 1)	Dummy
Socio-demographic characteristics		
Married	Married (yes = 1)	Dummy
Divorced	Divorced (yes = 1)	Dummy
German nationality	German (yes= 1)	Dummy
Number of children in household	Number of children in household	Metric
Age: 20-29 years (R)	20–29 years old (yes = 1)	Dummy
Age: 30-39 years	30–39 years old (yes = 1)	Dummy
Age: 40-52 years	40–52 years old (yes = 1)	Dummy
Level of Education		
Lower second. school/no degree (R)	Lower secondary school / no degree (yes = 1)	Dummy
Intermediate secondary school	Intermediate secondary school (yes = 1)	Dummy
Upper secondary school	Upper secondary school (yes = 1)	Dummy
Other graduation diploma	Other graduation diploma (yes = 1)	Dummy
In education (R)	In education (yes = 1)	Dummy
No vocational degree	No vocational degree (yes = 1)	Dummy
Vocational degree	Vocational degree below university (yes = 1)	Dummy
University degree	University degree (yes = 1)	Dummy
Sports participation		
Sports activity at least monthly	Sports activity at least monthly (yes = 1)	Dummy
Health indicators		
Satisfaction with health (0–10)	Satisfaction with health (scale 1–10)	Metric
Satisfaction with life (0–10)	Satisfaction with life (scale 1–10)	Metric
Number of doctor visits per year	Number of doctor visits per year	Metric
Physical health scale (0–100)	Physical component summary scale "physical health" (scale 0–100)	Metric
Mental health scale (0–100)	Mental component summary scale "mental health" (scale 0–100)	Metric
Social capital indicators		
Helping friends at least monthly	At least monthly (yes = 1)	Dummy
Voluntary work at least monthly	At least monthly (yes = 1)	Dummy
Civic involvement	Monthly or less often (yes = 1)	Dummy
Labor market indicators		
Household net income in EUR	Monthly household net income	Metric
Net earnings in EUR	Monthly net labor income	Metric
Gross earnings in EUR	Monthly gross labor income	Metric
Working time per week	Number of working hours per week	Metric
Hourly wage in EUR	Hourly wage	Metric
Fulltime employed	Full time employment (yes = 1)	Dummy

Table 5.B.2: Mean values and marginal effects of the selection models

	Men 20-52 years in 2001						Women 20-52 years in 2001					
	Mean in subsample			Marginal effects			Mean in subsample			Marginal effects		
Characteristics	High	Med	Low	H-L	H-M	M-L	High	Med	Low	H-L	H-M	M-L
(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)
Population size and density												
Population size ≥ 500k	0.44	0.38	0.32	R	R	R	0.46	0.41	0.32	R	R	R
Population size 100k-499k (L)	0.18	0.19	0.12	-0.20**	-0.53***	0.15	0.17	0.18	0.11	-0.04	-0.27***	0.05
Population size 50k-99k (M)	0.02	0.04	0.07	-0.09	-0.41	0.13	0.03	0.04	0.08	-0.002	-0.19	0.02
Population size 0-50k (S)	0.35	0.38	0.49	-0.11**	-0.42***	0.17**	0.34	0.37	0.49	-0.11*	-0.37***	0.10*
Population density	1.8	1.7	1.2	0.02	-0.14***	0.10***	1.8	1.8	1.3	0.02	-0.10***	0.06***
Population density x L	0.38	0.35	0.17	0.16***	0.22***	0.03	0.32	0.34	0.17	0.08*	0.10**	0.07**
Population density x M	0.04	0.07	0.10	-0.05	0.09	-0.06	0.04	0.07	0.11	-0.09*	-0.008	-0.03
Population density x S	0.29	0.22	0.17	0.08***	0.18***	-0.02	0.29	0.21	0.20	0.07**	0.17***	-0.008
Population characteristics												
Share of unemployed [1)]	7.1	8.9	9.2	-0.02***	0.02	-0.03***	7.2	8.9	9.4	-0.02***	0.01***	-0.03***
Share of graduates with HEEQ [1)]	24	25	23	0.003	-0.0003	0.005**	24	25	23	0.0009	-0.002	0.002
Share of foreigners	1.9	1.5	1.2	0.03	0.03	0.005	1.9	1.5	1.2	0.06**	0.02	0.04**
Weather conditions												
Average precipitation [2)]	2.1	2.0	2.0	0.33***	0.28***	0.06	2.1	2.0	2.0	0.29***	0.25***	0.02
Average sunshine duration [2)]	0.47	0.46	0.45	0.98***	0.12	0.80**	0.47	0.47	0.46	0.31	-0.51	0.34
Region												
North-Rhine Westphalia	0.22	0.20	0.21	R	R	R	0.25	0.20	0.20	R	R	R
Southern Germany	0.40	0.18	0.17	0.10**	0.20***	-0.05	0.38	0.19	0.16	0.12***	0.22***	-0.02
Northern Germany	0.04	0.14	0.20	-0.20***	-0.27***	-0.03	0.03	0.13	0.21	-0.29***	-0.31***	-0.09***
Western Germany	0.19	0.12	0.08	0.20***	0.14***	0.13**	0.19	0.12	0.09	0.16***	0.15***	0.07
Eastern Germany	0.14	0.35	0.34	0.13**	-0.17**	0.19***	0.15	0.36	0.34	0.17***	-0.16**	0.27***
Socio-demographic characteristics												
Married	0.65	0.63	0.64	-0.009	-0.01	-0.02	0.66	0.64	0.65	0.03	0.007	0.03
Divorced	0.05	0.06	0.06	-0.05	-0.02	-0.06	0.09	0.09	0.07	0.05	0.05	0.05
German nationality	0.87	0.92	0.94	-0.06	-0.02	-0.03	0.86	0.91	0.94	-0.11***	-0.02	-0.10**
Number of children in household	0.94	0.90	0.96	-0.02**	-0.007	-0.01	0.95	0.90	0.98	-0.02**	0.004	-0.02*
Age: 20–29 years	0.17	0.16	0.17	R	R	R	0.19	0.21	0.21	R	R	R
Age: 30–39 years	0.41	0.37	0.37	-0.01	-0.01	0.01	0.40	0.35	0.34	0.08***	0.06*	0.01
Age: 40–52 years	0.42	0.47	0.45	-0.04	-0.04	0.02	0.41	0.44	0.45	0.02	0.03	-0.02
Level of education												
Lower second. school/no degree	0.33	0.32	0.32	R	R	R	0.29	0.25	0.26	R	R	R
Intermediate secondary school	0.27	0.35	0.38	-0.01	0.02	-0.03	0.38	0.43	0.47	0.01	0.04	-0.03
Upper secondary school	0.30	0.27	0.26	-0.02	0.03	-0.06	0.24	0.25	0.22	0.08**	0.06	-0.004
Other graduation diploma	0.10	0.06	0.05	0.05	0.08	0.04	0.09	0.07	0.05	0.009	-0.006	0.008
In education	0.06	0.05	0.05	R	R	R	0.05	0.07	0.06	R	R	R
No vocational degree	0.12	0.11	0.11	-0.009	0.003	-0.05	0.19	0.12	0.14	0.01	0.10**	-0.10*
Vocational degree	0.71	0.74	0.76	-0.001	0.01	-0.05	0.67	0.72	0.73	0.004	0.02	-0.01
University degree	0.22	0.21	0.17	0.03	0.005	0.0004	0.16	0.20	0.19	-0.06*	-0.02	-0.04
Sports participation												
Sports activity at least monthly	0.40	0.37	0.37	-0.03	-0.02	-0.008	0.35	0.36	0.33	0.006	-0.005	0.01
Health indicators												
Satisfaction with health (0–10)	7.3	7.1	7.1	-0.007	0.005	-0.009	7.2	7.1	7.0	0.01*	0.006	0.007
Satisfaction with life (0–10)	7.2	7.0	7.0	0.01	0.005	0.005	7.2	7.1	7.1	-0.02**	-0.008	-0.01
Number of doctor visits per year	6.7	7.0	6.9	-0.0005	0.0002	-0.0001	9.9	10	10	-0.0002	0.0003	-0.0005
Physical health scale (0–100) [3)]	53	53	53	-0.0003	-0.0005	0.0009	53	52	52	-0.0009	0.0005	-0.0004
Mental health scale (0–100) [3)]	50	50	50	0.0002	0.0006	-0.0007	48	47	48	0.0006	0.001	-0.001
Social capital indicators												
Helping friends at least monthly	0.53	0.49	0.52	0.03*	0.06***	-0.01	0.46	0.43	0.43	0.04*	0.04**	0.02
Voluntary work at least monthly	0.19	0.17	0.21	-0.009	0.002	-0.01	0.13	0.13	0.13	-0.02	-0.03	0.02
Civic involvement	0.09	0.10	0.11	-0.05	-0.05	-0.0004	0.08	0.07	0.08	-0.04	0.06	-0.04
Labor market indicators												
Household net income in EUR	2733	2577	2592	0.000009	0.00001	-0.00001	2640	2568	2505	0.00002*	0.000005	0.00001
Gross earnings in EUR	2731	2424	2271	0.000007	0.00002	0.00001	1165	1165	1086	-0.00002	-0.00002	0.000006
Working time per week	41	39	38	0.0008	0.001	-0.0003	23	23	23	0.0006	0.002**	-0.001
Hourly wage in EUR	14	13	12	-0.0003	-0.004	-0.0005	8	8	8	0.003	-0.001	0.003
Fulltime employed	0.88	0.84	0.81	0.02	-0.009	0.04	0.39	0.39	0.38	0.019	0.006	-0.0008
Nr of observations, Pseudo R^2	985	1051	1009	22	13	7	1095	1150	1135	21	12	8

Note: All figures are based on the 2007 sample. For calculating the ATEs we needed to re-estimate all probit models for each year since we have an unbalanced panel with a decreasing number of observations due to panel attrition (results are available upon request). The 'high' ('med') ['low'] sample consists of individuals living in a municipality with high (medium) [low] levels of distance weighted per capita sports facility expenditures as defined in Chapter 5.2. The dependent variables in the probit models are dummies measuring one for the relatively higher expenditure level, i.e. for the H-M-model 'high', for the H-L-model 'high' and for the M-L-model 'med'. Independent variables are measured prior to the dependent variable if not stated otherwise, i.e. 1) measured between 2001 and 2006, 2) measured between 1981 and 2010, 3) measured in 2002. $***, **, *$ denote significance at the 1% , 5% , 10% levels. R denotes the reference category. A detailed description of all variables is provided in Table 5.B.1 in Appendix 5.B. Sources: RDC of the Federal Statistical Office and Statistical Offices of the Länder (2018a; 2018b) and SOEP as well as further data sources as discussed in detail in Chapters 5.2, 5.3 and in Appendix 5.A, own calculations.

Table 5.B.3: Common support statistics for the estimated matching models for men (with / without lagged SPE) and women (without lagged SPE)

			Men				**Women**	
			Without lagged SPE		With lagged SPE		Without lagged SPE	
Year	Model	Group	On support	Off support	On support	Off support	On support	Off support
2007	H-L	Untreated	1009	1	961	39	1135	0
		Treated	985	0	974	17	1095	29
	H-M	Untreated	1051	5	1035	17	1150	2
		Treated	985	6	974	0	1095	2
	M-L	Untreated	1009	14	961	24	1135	11
		Treated	1051	0	1035	4	1150	3
2008	H-L	Untreated	882	2	826	28	991	1
		Treated	820	22	699	149	913	39
	H-M	Untreated	885	22	871	12	981	1
		Treated	820	11	699	5	913	12
	M-L	Untreated	882	9	826	39	991	13
		Treated	885	0	871	24	981	1
2009	H-L	Untreated	830	0	795	38	955	0
		Treated	759	12	673	99	863	24
	H-M	Untreated	843	19	826	6	935	2
		Treated	759	11	673	10	863	4
	M-L	Untreated	830	14	795	11	955	11
		Treated	843	0	826	30	935	1
2010	H-L	Untreated	751	2	704	50	860	0
		Treated	677	10	599	90	770	10
	H-M	Untreated	745	23	731	11	840	0
		Treated	677	9	599	7	770	2
	M-L	Untreated	751	5	704	4	860	6
		Treated	745	0	731	26	840	0
2011	H-L	Untreated	722	0	664	43	827	1
		Treated	647	28	590	83	742	11
	H-M	Untreated	719	13	705	8	802	0
		Treated	647	6	590	8	742	1
	M-L	Untreated	722	14	664	29	827	5
		Treated	719	0	705	19	802	2
2012	H-L	Untreated	624	4	580	36	692	38
		Treated	544	19	478	88	615	8
	H-M	Untreated	607	24	597	11	679	28
		Treated	544	12	478	9	615	8
	M-L	Untreated	624	11	580	23	692	3
		Treated	607	1	597	24	679	2

Note: Common support statistics of the estimated matching models. The 'H' ('M') ['L'] refers to the corresponding models including individuals living in a municipality with high (medium) [low] levels of distance weighted per capita sports facility expenditures as defined in Chapter 5.2. Treated are those in the relatively higher expenditure level, i.e. for the H-M-model 'H', for the H-L-model 'H' and for the M-L-model 'M'. Sources: RDC of the Federal Statistical Office and Statistical Offices of the Länder (2018a; 2018b) and SOEP as well as further data sources as discussed in detail in Chapters 5.2, 5.3 and in Appendix 5.A, own calculations.

Table 5.B.4: Balancing statistics with the mean and median biases for the estimated matching models for men.

Year	Model		Mean	Med	Variables with % bias > 10 for the matched observations (M)
2007	H-L	UO	20.7	18.2	
		MO	8.2	6.7	gross earnings, hourly wage, average precipitation, share of foreigners, share of graduates with HEEQ, population size 100k–499k, population density, population density x L, upper secondary school, vocational degree, age: 30–39 years
	H-M	UO	14.3	9.2	
		MO	5.0	4.1	household net income, gross earnings, hourly wage, share of graduates with HEEQ
	M-L	UO	8.7	5.8	
		MO	4.5	3.7	population density, western Germany, age: 40–52 years
2008	H-L	UO	20.8	18.1	
		MO	10.5	9.3	satisfaction with health, satisfaction with life, Nr of doctor visits per year, mental health scale, civic involvement, average precipitation, average sunshine duration, share of foreigners, share of graduates with HEEQ, population density, population density x L, Married, Nr of children in household, upper secondary school, no vocational degree, vocational degree, western Germany, southern Germany, age: 30–39 years, age: 40–52 years
	H-M	UO	14.1	9.4	
		MO	4.2	3.2	share of graduates with HEEQ
	M-L	UO	9.6	6.2	
		MO	4.6	4.5	divorced, upper secondary school, university degree
2009	H-L	UO	20	16.8	
		MO	8.1	8.5	hourly wage, satisfaction with life, civic involvement, average precipitation, average sunshine duration, share of foreigners, population size 100k–499k, population density, population density x L, population density x S, upper secondary school, no vocational degree, vocational degree, western Germany, eastern Germany, age: 30–39 years, age: 40–52 years
	H-M	UO	13.4	9.2	
		MO	3.8	2.8	fulltime employed, mental health scale, share of graduates with HEEQ
	M-L	UO	9.4	5.5	
		MO	7.9	6.1	sports activity, helping friends, average precipitation, share of foreigners, share of graduates with HEEQ, population size 100k–499k, population density x L, divorced, German nationality, intermediate secondary school, upper secondary school, other graduation diploma, northern Germany, southern Germany, age: 30–39 years, age: 40–52 years
2010	H-L	UO	18.9	14.5	
		MO	7.9	8.4	Gross earnings, hourly wage, average precipitation, share of foreigners, population size 50k–99k, population density x M, population density x S, Nr of children in household, intermediate secondary school, upper secondary school, vocational degree, western Germany
	H-M	UO	12.5	7.8	
		MO	3.9	3.0	share of graduates with HEEQ
	M-L	UO	9.9	6.0	
		MO	5.2	4.8	share of foreigners, population size 100k–499k, population density
2011	H-L	UO	18.4	11.6	
		MO	7.4	5.3	household net income, civic involvement, average precipitation, share of foreigners, population density, population density x L, population density x S, number of children in household, upper secondary school, no vocational degree, vocational degree, Western Germany, Eastern Germany
	H-M	UO	13.2	7.2	
		MO	5.4	4.4	household net income, population size 0–50k, population size 100k–499k x S, Western Germany, Southern Germany
	M-L	UO	8.8	4.9	
		MO	5.2	4.0	gross earnings, hourly wage, average precipitation, share of foreigners, population size 100k–499k, population density, population density x L
2012	H-L	UO	18.8	12.8	
		MO	7.7	6.4	household net income, civic involvement, average precipitation, share of foreigners, population size 50k–99k, population density, population density x M, intermediate secondary school, upper secondary school, no vocational degree, vocational degree, western Germany, eastern Germany, age: 30–39 years, age: 40–52 years
	H-M	UO	12.2	6.8	
		MO	4.3	3.7	hourly wage, sports activity, other graduation diploma
	M-L	UO	9.7	5.5	
		MO	5.1	4.3	satisfaction with health, share of foreigners, population density, population density x S, Nr of children in household

Note: Balancing statistics with the mean and median (med) biases for the estimated matching models for men. The 'H' ('M') ['L'] refers to the corresponding models including individuals living in a municipality with high (medium) [low] levels of distance weighted per capita sports facility expenditures as defined in Chapter 5.2. Treated are those in the relatively higher expenditure level, i.e. for the H-M-model 'H', for the H-L-model 'H' and for the M-L-model 'M'. 'MO' refers to matched observations and 'UO' refers to unmatched observations. Sources: RDC of the Federal Statistical Office and Statistical Offices of the Länder (2018a; 2018b) and SOEP as well as further data sources as discussed in detail in Chapters 5.2, 5.3 and in Appendix 5.A, own calculations.

Table 5.B.5: Balancing statistics with the mean and median biases for the estimated matching models for men with lagged SPE.

Year	Model		Mean	Med	Variables with % bias > 10 for the matched observations (M)
2007	H-L	UO	22.1	17.6	
		MO	**14.4**	**13.4**	lagged SPE, gross earnings, hourly wage, fulltime employed, working time, satisfaction with life, civic involvement, average sunshine duration, share of foreigners, share of graduates with HEEQ, population size 0–50k, population density, population density x S, married, German nationality, Nr of children in household, upper secondary school, other graduation diploma, no vocational degree, vocational degree, university degree, western Germany, southern Germany, eastern Germany, age: 30–39 years
	H-M	UO	15.5	9.8	
		MO	**8.9**	**8.1**	lagged SPE, gross earnings, hourly wage, Nr of doctor visits per year, voluntary work, average precipitation, share of foreigners, share of unemployed
	M-L	UO	9.8	5.7	
		MO	**4.6**	**3.6**	lagged SPE, share of foreigners, share of graduates with HEEQ, population density, population size 0–50k, population density, population density x S, intermediate secondary school, eastern Germany
2008	H-L	UO	20.1	16.1	
		MO	**9.5**	**6.5**	lagged SPE, fulltime employed, satisfaction with health, mental health scale, average precipitation, share of foreigners, share of graduates with HEEQ, population size 50k-99k, population size 0-50k, population density, population density x M, German nationality, no vocational degree, vocational degree, western Germany, eastern Germany
	H-M	UO	13.3	9.6	
		MO	**4.4**	**3.3**	share of foreigners, , share of graduates with HEEQ, population size 0–50k, population density
	M-L	UO	9.9	5.7	
		MO	**5.5**	**4.5**	lagged SPE, Nr of doctor visits per year, share of foreigners, population size 0–50k, population density
2009	H-L	UO	20.7	15.8	
		MO	**13.9**	**10.1**	lagged SPE, gross earnings, hourly wage, fulltime employed, working time, satisfaction with life, Nr of doctor visits per year, physical health scale, civic involvement, average precipitation, share of foreigners, share of graduates with HEEQ, population size 100k–499k, population size 0–50k, population density, German nationality, Nr of children in household, upper secondary school, no vocational degree, university degree, western Germany, eastern Germany
	H-M	UO	13.4	7.9	
		MO	**8.9**	**8.1**	lagged SPE, gross earnings, hourly wage, working time, helping friends, share of foreigners, share of unemployed, share of graduates with HEEQ, population size 0–50k, population density, Nr of children in household, other graduation diploma, western Germany, eastern Germany, age: 30–39 years
	M-L	UO	9.9	5.3	
		MO	**4.7**	**2.8**	Nr of doctor visits per year, population size 0–50k, population density x S, married, Nr of children in household
2010	H-L	UO	19.5	11.6	
		MO	**12.4**	**10.3**	lagged SPE, gross earnings, hourly wage, fulltime employed, working time, satisfaction health, satisfaction life, physical health scale, average precipitation, average sunshine duration, share of foreigners, share of graduates with HEEQ, population size 100k–499k, population size 50k-99k, population size 0–50k, population density, married, German nationality, Nr of children in household, upper secondary school, other graduation diploma, western Germany, southern Germany, eastern Germany
	H-M	UO	12.6	7.5	
		MO	**6.5**	**5.4**	share of foreigners, share of unemployed, share of graduates with HEEQ, population size 0-50k, population density, university degree, eastern Germany
	M-L	UO	10.3	6.1	
		MO	**6.2**	**5.1**	lagged SPE, household net income, gross earnings, sports activity, share of foreigners, population size 50k–99k, population size 0–50k, population density, intermediate secondary school

Note: Table continues on next page.

Balancing statistics with the mean and median biases for the estimated matching models for men with lagged SPE - continued

Year	Model		Mean	Med	Variables with % bias > 10 for the matched observations (M)
2011	H-L	UO	19.7	11.5	
		MO	**12.8**	**11.8**	lagged SPE, household net income, gross earnings, hourly wage, fulltime employed, working time, satisfaction with health, satisfaction with life, physical health scale, sports activity, average sunshine duration, share of foreigners, share of graduates with HEEQ, population size 0–50k, population density, German nationality, upper secondary school, no vocational degree, university degree, western Germany, southern Germany, eastern Germany
	H-M	UO	13.7	8.4	
		MO	**8.2**	**5.5**	household net income, mental health scale, voluntary work, share of foreigners, share of unemployed, share of graduates with HEEQ, population size 0–50k, population density, vocational degree, university degree, eastern Germany
	M-L	UO	9.3	5.3	
		MO	**5.6**	**3.6**	lagged SPE, average precipitation, share of foreigners, western Germany, eastern Germany
2012	H-L	UO	19.4	12.5	
		MO	**12.1**	**10.8**	lagged SPE, gross earnings, hourly wage, satisfaction with health, satisfaction with life, Nr of doctor visits per year, physical health scale, helping friends, average precipitation, share of foreigners, share of unemployed, share of graduates with HEEQ, population size 100k–499k, population size 0–50k, population density, married, German nationality, upper secondary school, no vocational degree, university degree, western Germany, eastern Germany
	H-M	UO	11.8	6.5	
		MO	**5.8**	**3.8**	average precipitation, share of foreigners, share of unemployed, share of graduates with HEEQ, population size 100k–499k, population density, eastern Germany
	M-L	UO	9.8	6.2	
		MO	**5.2**	**4.8**	lagged SPE, civic involvement, population density, intermediate secondary school, university degree

Note: Balancing statistics with the mean and median biases for the estimated matching models for men with lagged SPE. The 'H' ('M') ['L'] refers to the corresponding models including individuals living in a municipality with high (medium) [low] levels of distance weighted per capita sports facility expenditures as defined in Chapter 5.2. Treated are those in the relatively higher expenditure level, i.e. for the H-M-model 'H', for the H-L-model 'H' and for the M-L-model 'M'. 'MO' refers to matched observations and 'UO' refers to unmatched observations. Sources: RDC of the Federal Statistical Office and Statistical Offices of the Länder (2018a; 2018b) and SOEP as well as further data sources as discussed in detail in Chapters 5.2, 5.3 and in Appendix 5.A, own calculations.

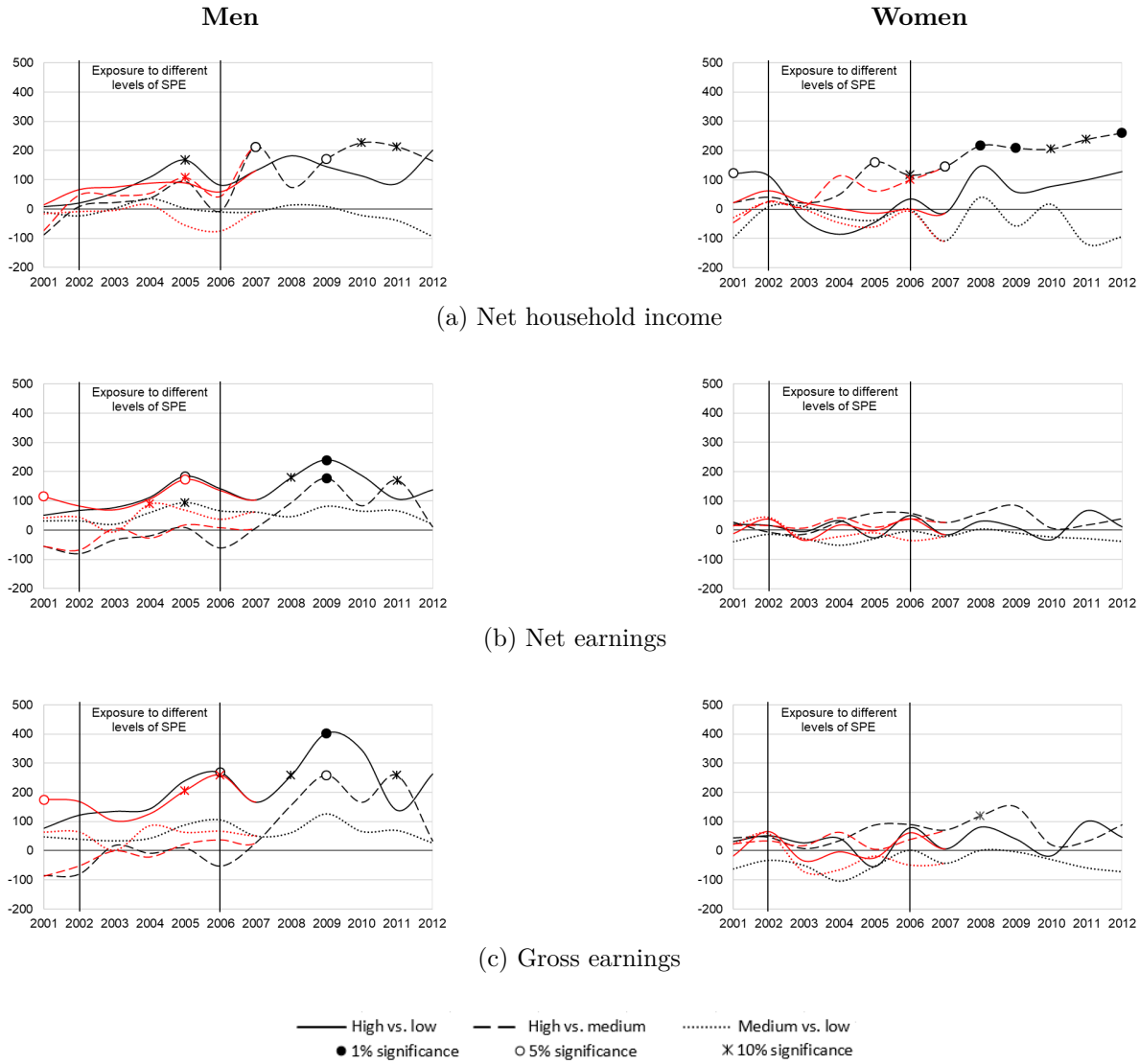
Table 5.B.6: Balancing statistics with the mean and median biases for the estimated matching models for women.

Year	Model		Mean	Med	Variables with % bias > 10 for the matched observations (M)
2007	H-L	UO	17.9	11.8	
		MO	**5.9**	**4.6**	average precipitation, share of foreigners, population size 100k–499k, population density, population density x L
	H-M	UO	11.4	6.2	
		MO	**3.4**	**2.2**	share of graduates with HEEQ
	M-L	UO	9.0	6.3	
		MO	**5.0**	**4.0**	household net income, population density x L
2008	H-L	UO	18.0	11.8	average precipitation, share of graduates with HEEQ, population size 100k–499k, population density x L, population density x S,
		MO	**6.5**	**4.6**	divorced, no vocational degree, vocational degree
	H-M	UO	11.8	7.4	
		MO	**4.1**	**3.1**	share of graduates with HEEQ
	M-L	UO	9.0	6.7	
		MO	**3.3**	**2.3**	intermediate secondary school, age: 30–39 years
2009	H-L	UO	18.4	12.1	average precipitation, average sunshine duration, share of foreigners, share of graduates with HEEQ, population density, divorced, German nationality
		MO	**6.2**	**4.7**	
	H-M	UO	11.8	6.1	
		MO	**4.0**	**3.1**	working time, share of foreigners, share of graduates with HEEQ, population density
	M-L	UO	9.5	5.3	
		MO	**3.9**	**2.5**	population size 100k–499k, population density x L, divorced
2010	H-L	UO	19.0	12.3	
		MO	**6.4**	**5.8**	average precipitation, divorced, German nationality, no vocational degree, population density x L
	H-M	UO	11.8	5.8	
		MO	**4.9**	**4.3**	fulltime employed, share of foreigners, share of graduates with HEEQ, population density, no vocational degree
	M-L	UO	9.9	6.4	
		MO	**4.6**	**3.5**	sports activity
2011	H-L	UO	18.0	9.9	working time, average precipitation, share of foreigners, share of graduates with HEEQ, population size 0–50k, population density,
		MO	**6.6**	**5.2**	divorced, German Nationality
	H-M	UO	11.4	5.6	
		MO	**4.7**	**3.6**	mental health scale, share of foreigners, share of graduates with HEEQ, German nationality, no vocational degree
	M-L	UO	9.3	4.8	
		MO	**5.0**	**2.8**	share of foreigners, population size 100k-499k, population density, population density x L, married, divorced, eastern Germany, age: 30–39 years
2012	H-L	UO	17.1	11.6	
		MO	**7.4**	**6.2**	working time, satisfaction with health, helping friends, civic involvement, share of foreigners, population density x S, married, divorced, German nationality, other graduation diploma, no vocational degree, southern Germany
	H-M	UO	10.8	5.8	
		MO	**5.5**	**4.5**	average sunshine duration, share of foreigners, share of graduates with HEEQ, population density, western Germany, southern Germany
	M-L	UO	9.1	5.9	
		MO	**5.6**	**5.0**	sports activity, helping friends, civic involvement, population size 100k–499k, population density x L, married, southern Germany, age: 30–39 years

Note: Balancing statistics with the mean and median (med) biases for the estimated matching models for women. The 'H' ('M') ['L'] refers to the corresponding models including individuals living in a municipality with high (medium) [low] levels of distance weighted per capita sports facility expenditures as defined in Chapter 5.2. Treated are those in the relatively higher expenditure level, i.e. for the H-M-model 'H', for the H-L-model 'H' and for the M-L-model 'M'. 'MO' refers to matched observations and 'UO' refers to unmatched observations. Sources: RDC of the Federal Statistical Office and Statistical Offices of the Länder (2018a; 2018b) and SOEP as well as further data sources as discussed in detail in Chapters 5.2, 5.3 and in Appendix 5.A, own calculations.

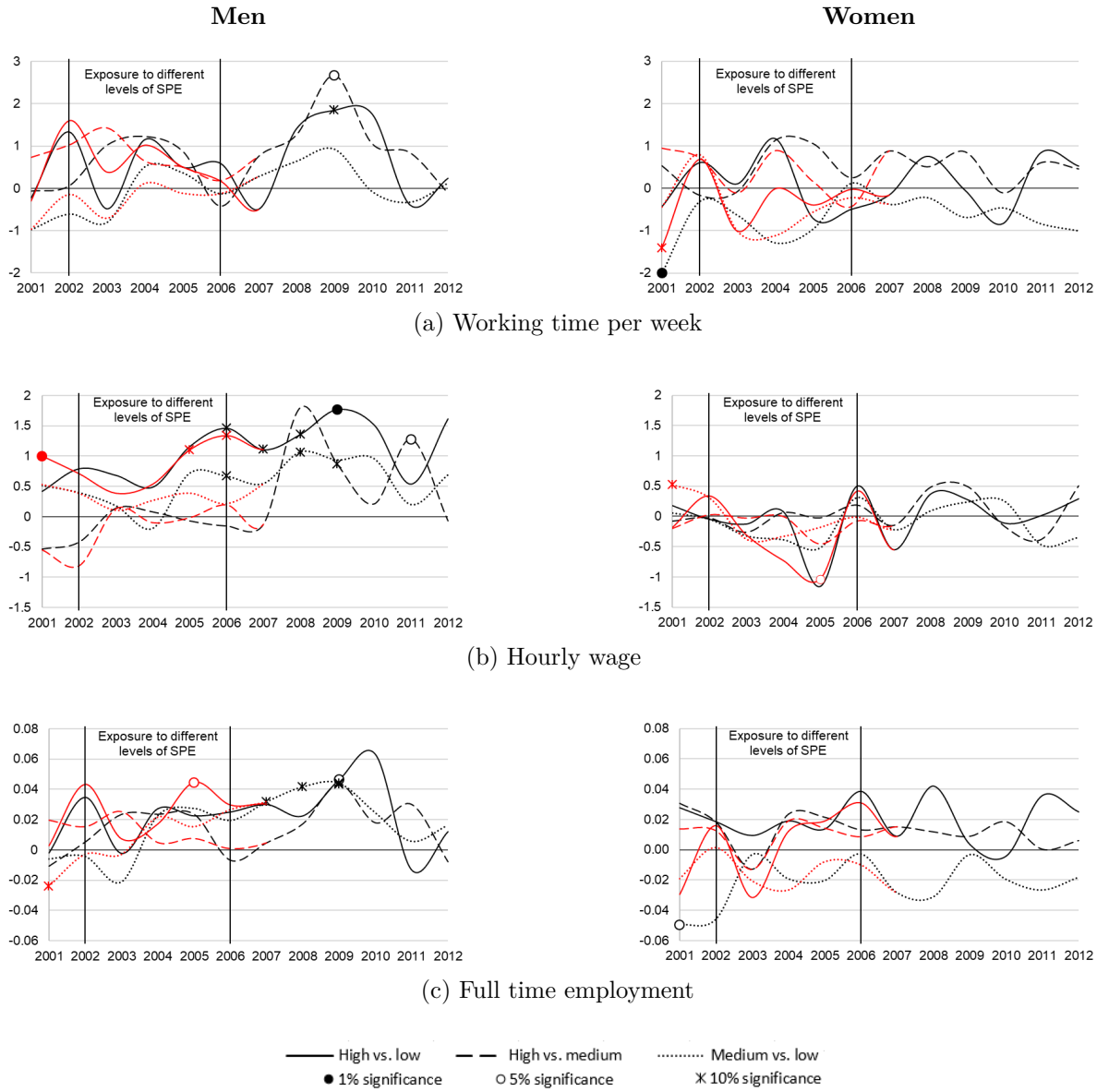
5.C Robustness check

Figure 5.C.1: Effects of different levels of SPE on household income and earnings during the exposure period including (main specification, black lines) / excluding (red lines) any individuals who were still in the education system in 2001-2006.



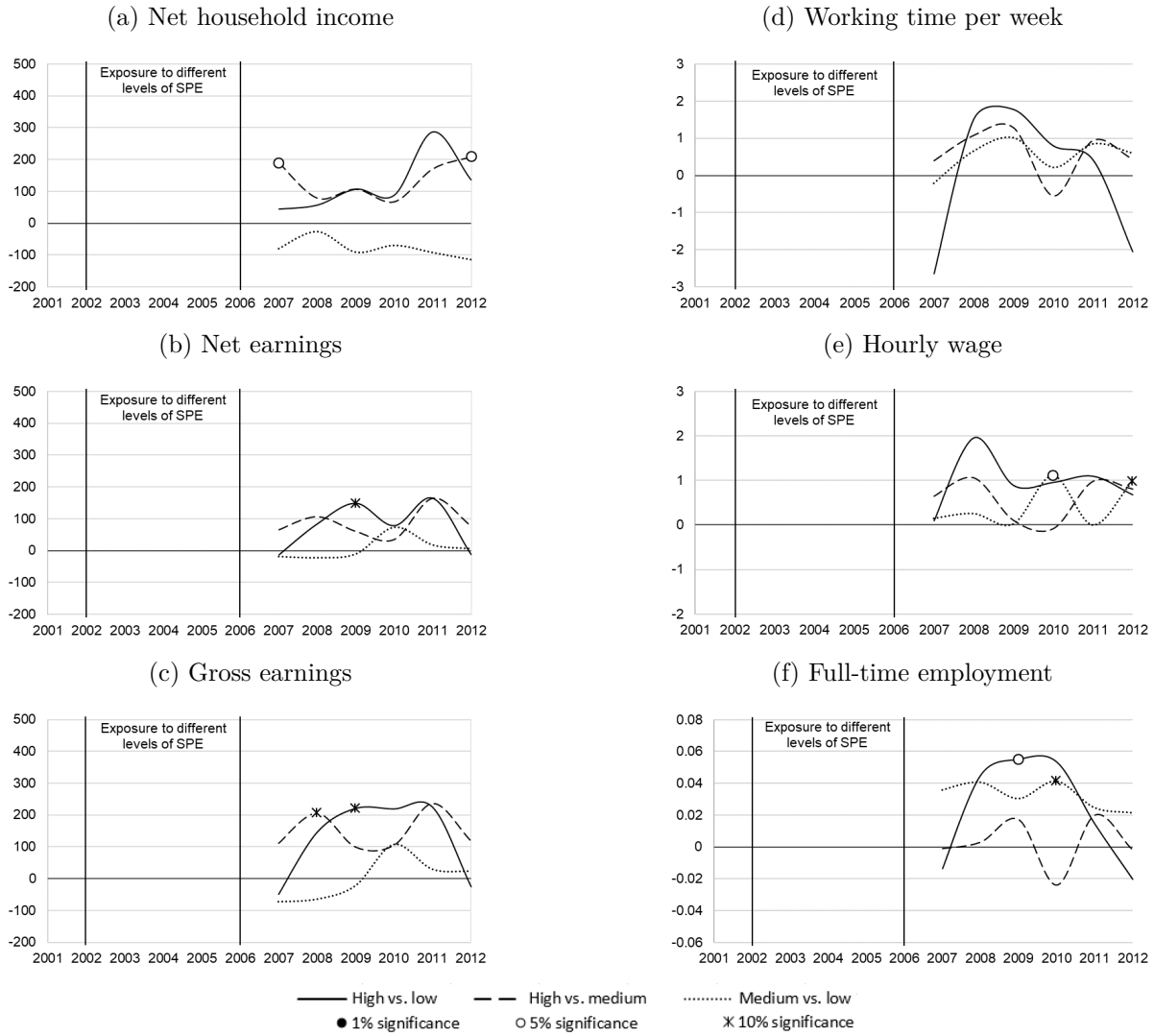
Note: Radius matching results for men and women during the exposure period including (main specification, black lines) / excluding (red lines) any individuals who were still in the education system in 2001-2006, i.e. trainees, students, civilian servants, conscripts or else. Displayed are Average Treatment Effects (ATEs) of 'high vs. low', 'high vs. medium' and 'medium vs. low' levels of distance weighted per capita sports facility expenditures (as defined in Chapter 5.2) on household income and earnings. All 2001 ATEs are calculated as mean differences between matched treated and untreated observations coming from the balancing tests of the 2002 estimations. The effects for net earnings in 2001 are approximated by multiplying the effects for gross earnings with the samples' average net-to-gross earnings ratio, i.e. 65.6% (63.5%) for men (women). Sources: RDC of the Federal Statistical Office and Statistical Offices of the Länder (2018a; 2018b) and SOEP as well as further data sources as discussed in detail in Chapters 5.2, 5.3 and in Appendix 5.A, own calculations.

Figure 5.C.2: Effects of different levels of SPE on working time, wages and employment status during the exposure period including (main specification, black lines) / excluding (red lines) any individuals who were still in the education system in 2001-2006.



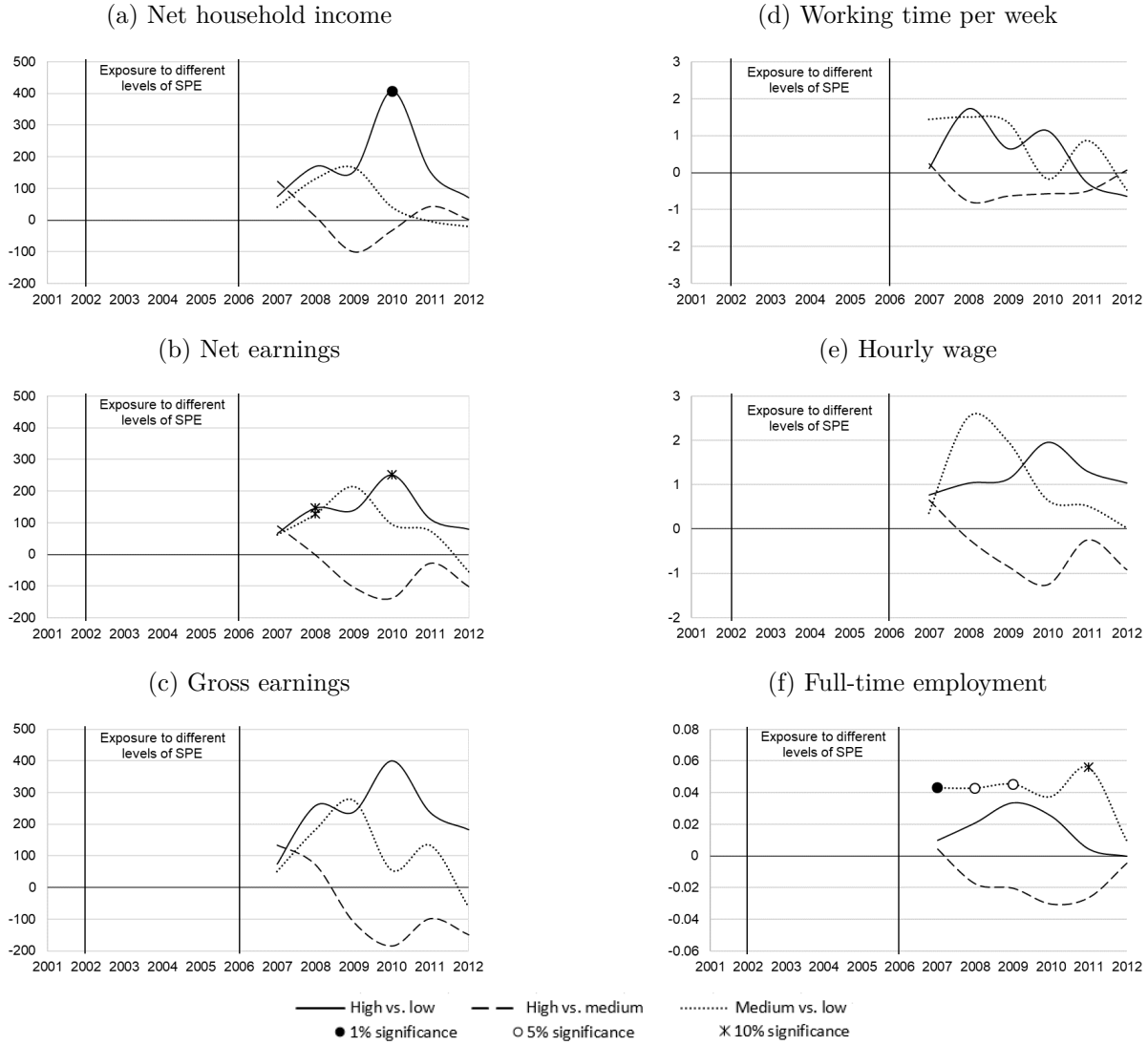
Note: Radius matching results for men and women during the exposure period including (main specification, black lines) / excluding (red lines) any individuals who were still in the education system in 2001-2006, i.e. trainees, students, civilian servants, conscripts or else. Displayed are Average Treatment Effects (ATEs) of 'high vs. low', 'high vs. medium' and 'medium vs. low' levels of distance weighted per capita sports facility expenditures (as defined in Chapter 5.2) on hourly wages, employment status and working time per week in hours. All 2001 ATEs are calculated as mean differences between matched treated and untreated observations coming from the balancing tests of the 2002 estimations. The effects for net earnings in 2001 are approximated by multiplying the effects for gross earnings with the samples' average net-to-gross earnings ratio, i.e. 65.6% (63.5%) for men (women). Sources: RDC of the Federal Statistical Office and Statistical Offices of the Länder (2018a; 2018b) and SOEP as well as further data sources as discussed in detail in Chapters 5.2, 5.3 and in Appendix 5.A, own calculations.

Figure 5.C.3: Effects of different levels of SPE on household income, earnings, working time, wages and full time employment for men when including lagged SPE as additional control.



Note: Radius matching results for men when including lagged SPE as additional control. Displayed are Average Treatment Effects (ATEs) of 'high vs. low', 'high vs. medium' and 'medium vs. low' levels of distance weighted per capita sports facility expenditures (as defined in Chapter 5.2) on household income, earnings, working time, wages and full time employment. Due to computational constraints caused by the infrastructure at the DIW, we were not able to estimate these models also for the pre-exposure and exposure periods. Sources: RDC of the Federal Statistical Office and Statistical Offices of the Länder (2018a; 2018b) and SOEP as well as further data sources as discussed in detail in Chapters 5.2, 5.3 and in Appendix 5.A, own calculations.

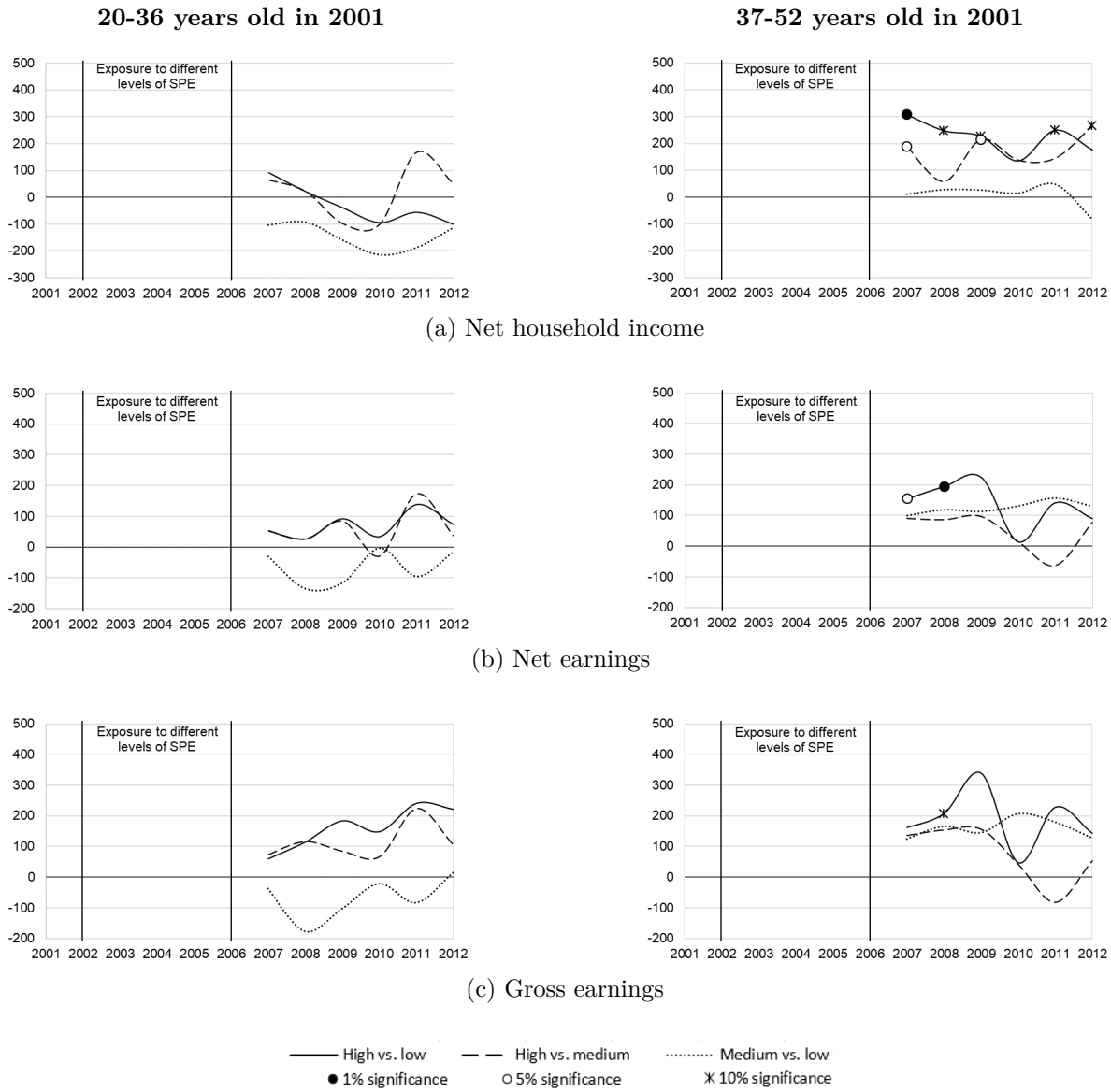
Figure 5.C.4: Effects of different levels of SPE on household income, earnings, working time, wages and full time employment for men when using the pure SPE measure.



Note: Radius matching results for men when using the original 'pure' SPE measure. Displayed are Average Treatment Effects (ATEs) of 'high vs. low', 'high vs. medium' and 'medium vs. low' levels of per capita sports facility expenditures on household income, earnings, working time, wages and full time employment. Due to computational constraints caused by the infrastructure at the DIW, we were not able to estimate these models also for the pre-exposure and exposure periods. Sources: RDC of the Federal Statistical Office and Statistical Offices of the Länder (2018a; 2018b) and SOEP as well as further data sources as discussed in detail in Chapters 5.2, 5.3 and in Appendix 5.A, own calculations.

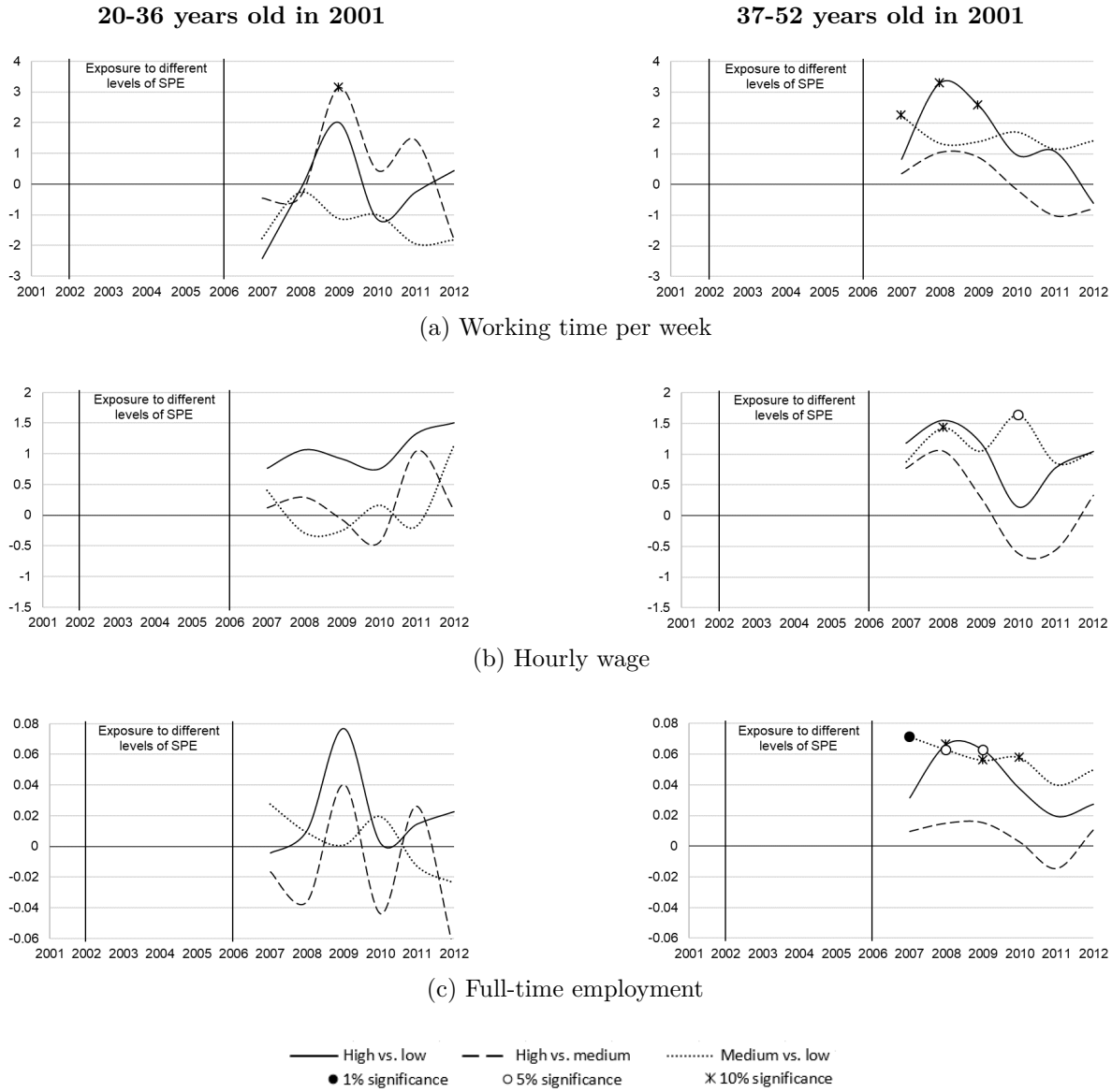
5.D Effect heterogeneity

Figure 5.D.1: Effects of different levels of SPE on earnings for men 20-36 / 37-52 years old in 2001.



Note: Radius matching results for men 20-36 / 37-52 years old in 2001. Displayed are Average Treatment Effects (ATEs) of 'high vs. low', 'high vs. medium' and 'medium vs. low' levels of distance weighted per capita sports facility expenditures (as defined in Chapter 5.2) on household income and earnings. Due to computational constraints caused by the infrastructure at the DIW, we were not able to estimate these models also for the pre-exposure and exposure periods. Sources: RDC of the Federal Statistical Office and Statistical Offices of the Länder (2018a; 2018b) and SOEP as well as further data sources as discussed in detail in Chapters 5.2, 5.3 and in Appendix 5.A, own calculations.

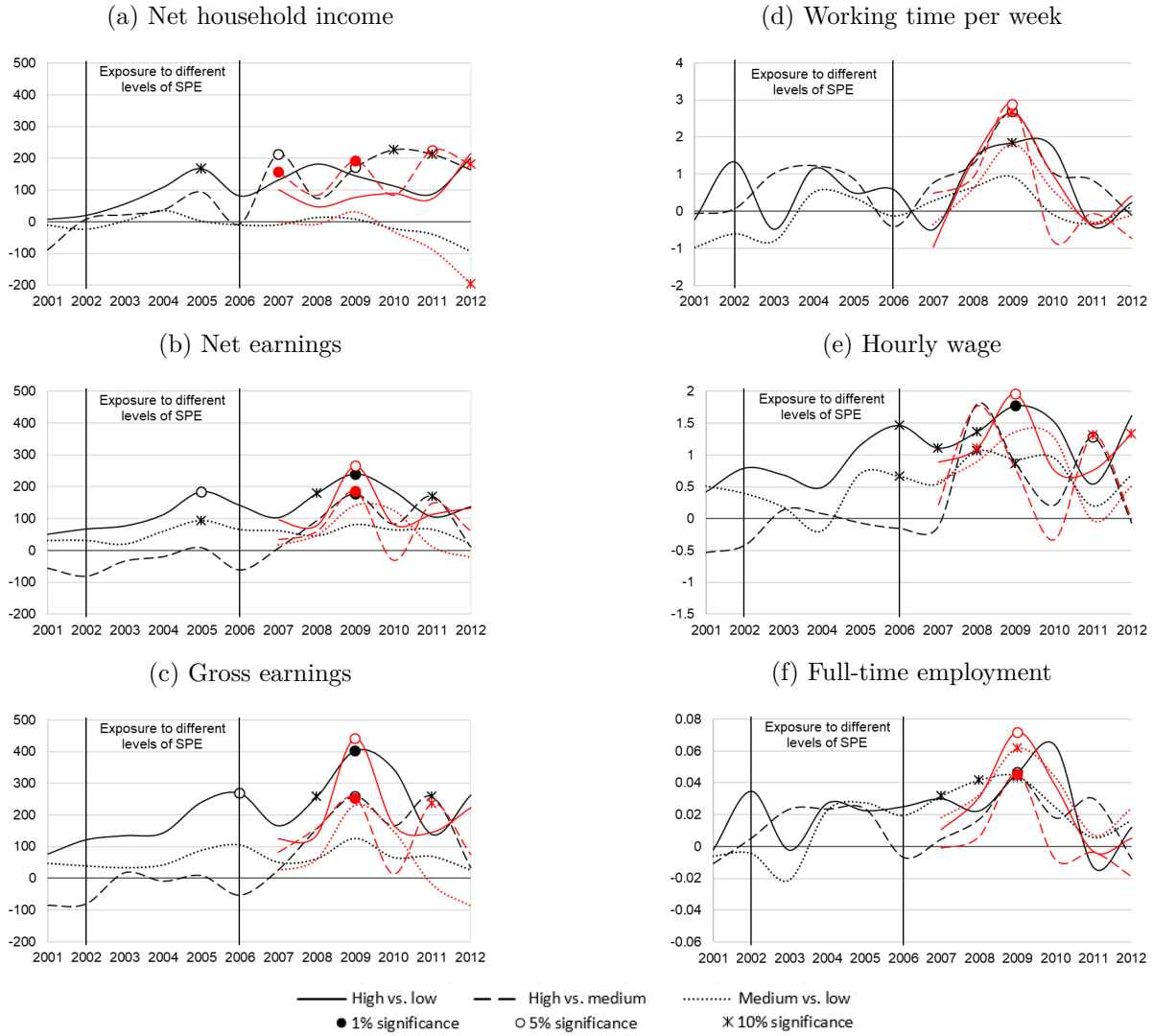
Figure 5.D.2: Effects of different levels of SPE on working time, wage and employment status for men 20-36 / 37-52 years old in 2001.



Note: Radius matching results for men for men 20-36 / 37-52 years old in 2001. Displayed are Average Treatment Effects (ATEs) of 'high vs. low', 'high vs. medium' and 'medium vs. low' levels of distance weighted per capita sports facility expenditures (as defined in Chapter 5.2) on hourly wages, employment status and working time per week in hours. Due to computational constraints caused by the infrastructure at the DIW, we were not able to estimate these models also for the pre-exposure and exposure periods. Sources: RDC of the Federal Statistical Office and Statistical Offices of the Länder (2018a; 2018b) and SOEP as well as further data sources as discussed in detail in Chapters 5.2, 5.3 and in Appendix 5.A, own calculations.

5.E Exploring the channels

Figure 5.E.1: Effects of different levels of SPE on household income, earnings, working time, wages and full time employment for men during the post-exposure period excluding (main specification, black lines) / including (red lines) various health measures from the corresponding year as additional covariates in the selection process.



Note: Radius matching results for men during the post-exposure period excluding (main specification, black lines) / including (red lines) satisfaction with health, satisfaction with life, the number of doctor visits per year as well as the physical component (PCS) and the mental component summary scales (MCS) from the corresponding year as additional covariates in the selection process. Displayed are Average Treatment Effects (ATEs) of 'high vs. low', 'high vs. medium' and 'medium vs. low' levels of distance weighted per capita sports facility expenditures (as defined in Chapter 5.2) on household income, earnings, working time, wages and full time employment. All 2001 ATEs are calculated as mean differences between matched treated and untreated observations coming from the balancing tests of the 2002 estimations. The effects for net earnings in 2001 are approximated by multiplying the effects for gross earnings with the samples' average net-to-gross earnings ratio, i.e. 65.6% (63.5%) for men (women). Sources: RDC of the Federal Statistical Office and Statistical Offices of the Länder (2018a; 2018b) and SOEP as well as further data sources as discussed in detail in Chapters 5.2, 5.3 and in Appendix 5.A, own calculations.

Table 5.E.1: Average values of the SPE measures and different indicators of sports facility provision in the 16 states (Länder) in Germany

	SPE measures in Euro [1]		**Nr sports facilities** [2]	**Area covered by pools in m^2 per 1,000 inhabitants** [3]		**Estimated need for refurbishment in %** [4]		
State in Germany (Land)	Pure	Radius	Total	Outdoor	Indoor	Fields	Gyms	Pools
South								
Baden-Wuerttemberg	52.0	48.8	18,556	47.5	14.7	*(15.0)* [5]	*(11.2)* [5]	*(35.9)* [5]
Bavaria	17.3	21.7	22,027	59.7	14.8	26.1	36	46.9
North								
Lower-Saxony	10.7	13.9	14,723	54.3	16.2	16.3	20.8	14.7
Schleswig Holstein	9.3	15.2	4,408	35.2	15.6	32.0	36.6	52.1
Bremen	25.3	22.2	741	34.1	12.4	57.8	66.4	65.2
Hamburg	7.4	12.1	1,852	11.5	8.1	*(5.6)* [5]	*(16.6)* [5]	30.8
West								
Hessen	31.9	32.9	10,189	54.7	15.5	35.9	40.8	54.8
Rhineland-Palatine	7.5	15.0	7,401	54.4	12.4	35.1	36.5	50.6
Saarland	44.1	42.4	1,659	73.4	18.8	51.1	58.6	70.9
North Rhine-Westphalia	27.8	27.2	21,531	31.7	15.2	34.0	36.8	55.8
East								
Brandenburg	10.0	12.2	3,978	31.7	7.2	70.8	56.7	56.5
Mecklenburg-West Pomerania	8.6	16.9	2,719	28.8	4.5	76.6	69.7	59.4
Saxony	23.6	25.7	6,096	121.7	8.2	71.8	70.2	62.1
Saxony-Anhalt	13.2	20.9	4,036	77.6	4.9	*(57.3)* [5]	*(53.5)* [5]	52.9
Thuringia	12.2	17.6	3,943	106.1	8.7	74.4	73.9	70.2
Berlin	29.1	29.1	3,095	11.6	10.2	52	56.1	76.8
Germany	16.9	21.4	126,954	52.1	13.3	36.7	41.8	49.1
Correlation with pure SPE measure [6]			27.9	14.9	47.2*			
Correlation with the SPE radius measure [6]			27.6	8.3	51.6**			

Note: 1) Pure and distance weighted per capita sports facility expenditure measures (mean of 2002-2006) as defined in Chapter 5.2 (Sources: RDC of the Federal Statistical Office and Statistical Offices of the Länder (2018a; 2018b), own calculations.).
2) Total number of sports facilities as of 1st July 2000 (Source: Sports Minister Conference (2003).
3) Areas covered by indoor and outdoor pools in m^2 per 1,000 inhabitants as of 1st July 2000 (Source: Sports Minister Conference (2003).
4) Estimated need for restoration in % of sports fields, gyms and pools as of 1st July 2000, excluding 400 meter-tracks, sport fields $< 5,000m^2$, tennis courts and facilities for shooting sports (Source: Sports Minister Conference (2003).
5) Unreliable estimates due to missing values $> 20\%$ (source: Sports Minister Conference (2003).
6) Correlations between the corresponding SPE measure and the indicators of sports facility provision based on information for N=16 states ('Länder'), $***, **, *$ denote significance at the 1% , 5% , 10% levels. (Sources: RDC of the Federal Statistical Office and Statistical Offices of the Länder (2018a; 2018b), Sports Minister Conference (2003), own calculations.)

Table 5.E.2: Comparing the effects for men between our main specification (ATE) and the main specification including various health measures from the corresponding year as additional covariates in the selection process (ATE_{Health}).

Outcome	Model	2007 ATE	2007 ATE_{Health}	2008 ATE	2008 ATE_{Health}	2009 ATE	2009 ATE_{Health}	2010 ATE	2010 ATE_{Health}	2011 ATE	2011 ATE_{Health}	2012 ATE	2012 ATE_{Health}	Average change[1]	Average change[2]
Household net income in Euros	H-L	131	100	182	46	145	76	113	89	87	72	200	214		
	H-M	**213****	**157*****	74	84	**171****	**191*****	**227***	**84**	**213***	**225****	**164**	**181***	-13.2%	-16.9%
	M-L	-9	0	14	-7	9	32	-21	-31	-39	-87	-93	-195*		
Net earnings in Euros	H-L	103	96	**179***	**79**	**240*****	**265****	186	82	106	114	137	134	-18.9%	-23.4%
	H-M	7	33	93	60	**177*****	**185*****	82	-32	**170***	**148**	11	61	-15.8%	-8.2%
	M-L	62	18	46	49	82	143	65	126	66	11	20	-22		
Gross earnings in Euros	H-L	166	124	**258***	**137**	**403*****	**440****	345	168	138	145	263	223	-21.4%	-15.9%
	H-M	27	82	156	158	**259****	**253*****	166	14	**258***	**236***	33	76	-9.0%	-9.6%
	M-L	51	25	62	59	127	230	66	148	70	-19	26	-87		
Working time per week in hours	H-L	-0.5	-1.0	1.4	1.4	1.8*	2.7*	1.8	1.0	-0.4	-0.3	0.2	0.4		
	H-M	0.8	0.5	1.3	0.9	2.7**	2.9**	1.1	-0.8	0.8	-0.1	-0.1	-0.7		
	M-L	0.3	-0.4	0.6	0.6	0.9	1.8	-0.1	0.6	-0.3	-0.3	0.1	-0.1		
Hourly wage in Euros	H-L	**1.1***	**0.9**	**1.4***	**1.1***	**1.8*****	**2.0****	1.5	0.8	0.5	0.8	**1.6**	**1.3****	-14.4%	-11.2%
	H-M	-0.1	0.2	1.8	1.8	**0.9***	**0.8**	0.2	-0.3	**1.3****	**1.3****	-0.1	0.0	-5.8%	-4.6%
	M-L	0.5	0.6	1.1*	0.9	0.9	1.4	1.0	1.3	0.2	0.0	0.7	0.5		
Full time employment in percent	H-L	3.0	1.1	2.3	3.1	4.7**	7.2**	6.3	3.8	-1.3	-0.3	1.2	0.5		
	H-M	0.5	-0.1	1.7	0.6	4.3*	4.5***	1.8	-0.8	3.0	-0.4	-0.8	-1.9		
	M-L	**3.2***	**1.8**	**4.2***	**3.3**	**4.4***	**6.2***	2.5	4.4	0.6	0.7	1.6	2.4	13.8%	-6.2%

*Note:*Displayed are Average Treatment Effects (ATEs) of 'high vs. low' (H-L), 'high vs. medium' (H-M) and 'medium vs. low' (M-L) levels of distance weighted per capita sports facility expenditures (as defined in Chapter 5.2) household income, earnings, working time, wages and full time employment. ATE = results derived from our main specification, ATE_{Health} = results derived from our main specification including satisfaction with health, satisfaction with life, the number of doctor visits per year as well as the physical component (PCS) and the mental component summary scales (MCS) from the corresponding year as additional covariates in the selection model (see Table 5.B.1 in the Appendix for further information on the variable definition / measurement as well as Andersen et al. (2007) for a detailed description of the PCS and MCS). $***, **, *$ denote significance at the 1% , 5% , 10% levels. Average change[1] = average change (in percent) between ATE and ATE_{Health} calculated for all years 2007-2012 (only calculated if either ATE and/or ATE_{Health} are significant for at least two years, displayed in bold letters), Average change[2] = average change (in percent) between ATE and ATE_{Health} calculated only for years with significant estimates (in bold letters) (only calculated if either ATE and/or ATE_{Health} are significant for at least two years, displayed in bold letters). Sources: RDC of the Federal Statistical Office and Statistical Offices of the Länder (2018a; 2018b) and SOEP as well as further data sources as discussed in detail in Chapters 5.2, 5.3 and in Appendix 5.A, own calculations.

References

Aaronson, D., Lange, F., & Mazumder, B. (2014). Fertility Transitions along the Extensive and Intensive Margins. *American Economic Review*, *104*(11), 3701–24.

Abramitzky, R., Delavande, A., & Vasconcelos, L. (2011). Marrying Up: The Role of Sex Ratio in Assortative Matching. *American Economic Journal: Applied Economics*, *3*(3), 124–157.

Acemoglu, D., Autor, D. H., & Lyle, D. (2004). Women, War, and Wages: The Effect of Female Labor Supply on the Wage Structure at Midcentury. *Journal of Political Economy*, *112*(3), 497–551.

Ahlfeldt, G. M., & Kavetsos, G. (2014). Form or function?: The effect of new sports stadia on property prices in London. *Journal of the Royal Statistical Society: Series A (Statistics in Society)*, *177*(1), 169–190.

Aigner, D. J., & Cain, G. G. (1977). Statistical Theories of Discrimination in Labor Markets. *Industrial and Labor Relations Review*, *30*(2), 175–187.

Akbulut-Yuksel, M. (2014). Children of War: The Long-Run Effects of Large-Scale Physical Destruction and Warfare on Children. *Journal of Human Resources*, *49*(3), 634–662.

Akerlof, G. A., & Kranton, R. E. (2000). Economics and Identity. *The Quarterly Journal of Economics*, *115*(3), 715–753.

Alba, R. D., & Logan, J. R. (1991). Variations on two themes: Racial and Ethnic Patterns in the Attainment of Suburban Residence. *Demography*, *28*(3), 431–453.

Andersen, H. H., Mühlbacher, A., Nübling, M., Schupp, J., & Wagner, G. G. (2007). Computation of Standard Values for Physical and Mental Health Scale Scores Using the SOEP Version of SF12v2. *Schmollers Jahrbuch*, *127*(1), 171–182.

Angrist, J. D. (2002). How Do Sex Ratios Affect Marriage and Labor Markets? Evidence from America's Second Generation. *The Quarterly Journal of Economics*, *117*(3), 997–1038.

Angrist, J. D., Imbens, G. W., & Rubin, D. B. (1996). Identification of Causal Effects using Instrumental Variables. *Journal of the American Statistical Association*,

91(434), 444–455.

Angrist, J. D., & Krueger, A. B. (1992). The Effect of Age at School Entry on Educational Attainment: An Application of Instrumental Variables with Moments from Two Samples. *Journal of the American Statistical Association*, *87*(418), 328–336.

Angrist, J. D., & Krueger, A. B. (1995). Split-sample instrumental variables estimates of the return to schooling. *Journal of Business & Economic Statistics*, *13*(2), 225–235.

Anukriti, S. (2018). Financial Incentives and the Fertility-Sex Ratio Trade-Off. *American Economic Journal: Applied Economics*, *10*(2), 27–57.

Arellano, M., & Meghir, C. (1992). Female Labour Supply and On-the-Job Search: An Empirical Model Estimated Using Complementary Data Sets. *The Review of Economic Studies*, *59*(3), 537–559.

Ashraf, N., Field, E., & Lee, J. (2014). Household Bargaining and Excess Fertility: An Experimental Study in Zambia. *American Economic Review*, *104*(7), 2210–37.

Attané, I. (2006). The Demographic Impact of a Female Deficit in China, 2000–2050. *Population and Development Review*, *32*(4), 755–770.

Ausschuss der deutschen Statistiker für die Volks– und Berufszählung 1946. (1949a). *Volks- und Berufszählung vom 29. Oktober 1946 in den vier Besatzungszonen und Gross-Berlin: Deutsches Gemeindeverzeichnis*. Berlin and Munich: Duncker and Humblot.

Ausschuss der deutschen Statistiker für die Volks– und Berufszählung 1946. (1949b). *Volks- und Berufszählung vom 29. Oktober 1946 in den vier Besatzungszonen und Gross-Berlin: Volkszählung Tabellenteil*. Berlin and Munich: Duncker and Humblot.

Ausschuss der deutschen Statistiker für die Volks– und Berufszählung 1946. (1951). *Volks- und Berufszählung vom 29. Oktober 1946 in den vier Besatzungszonen und Gross-Berlin: Volkszählung Textteil*. Berlin and Munich: Duncker and Humblot.

Ausschuss der deutschen Statistiker für die Volks– und Berufszählung 1946. (1953). *Volks- und Berufszählung vom 29. Oktober 1946 in den vier Besatzungszonen und Gross-Berlin: Berufszählung Textteil*. Berlin and Munich: Duncker and Humblot.

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.

Baltrunaite, A., Casarico, A., Profeta, P., & Savio, G. (2016). *Let the Voters Choose Women.* CESifo Working Paper No. 5693.

Baudin, T., De la Croix, D., & Gobbi, P. E. (2015). Fertility and Childlessness in the United States. *American Economic Review*, *105*(6), 1852–82.

Bauer, M., Blattman, C., Chytilová, J., Henrich, J., Miguel, E., & Mitts, T. (2016). Can War Foster Cooperation? *Journal of Economic Perspectives*, *30*(3), 249–274.

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

Beaman, L., Chattopadhyay, R., Duflo, E., Pande, R., & Topalova, P. (2009). Powerful Women: Does Exposure Reduce Bias? *The Quarterly Journal of Economics*, *124*(4), 1497–1540.

Beaman, L., Duflo, E., Pande, R., & Topalova, P. (2012). Female Leadership Raises Aspirations and Educational Attainment for Girls: A Policy Experiment in India. *Science*, *335*(6068), 582–586.

Becker, G. S. (1973). A Theory of Marriage: Part I. *Journal of Political Economy*, *81*(4), 813–846.

Becker, G. S. (1974). A Theory of Marriage. In T. W. Schultz (Ed.), *Economics of the Family: Marriage, Children, and Human Capital* (pp. 299–351). University of Chicago Press.

Becker, G. S. (1991). *A Treatise on the Family (Revised and enlarged edition).* Cambridge, Massachusetts: Harvard University Press.

Bernhardt, A., Field, E., Pande, R., Rigol, N., Schaner, S., & Troyer-Moore, C. (2018). Male Social Status and Women's Work. In *AEA Papers and Proceedings* (Vol. 108, pp. 363–67).

Bertrand, M. (2011). New Perspectives on Gender. In O. Ashenfelter & D. Card (Eds.), *Handbook of Labor Economics* (Vol. 4B, pp. 1543–1590). Elsevier.

Bertrand, M., Duflo, E., & Mullainathan, S. (2004). How Much Should We Trust

Differences-In-Differences Estimates? *The Quarterly Journal of Economics*, *119*(1), 249–275.

Bethmann, D., & Kvasnicka, M. (2013). World War II, Missing Men and Out of Wedlock Childbearing. *The Economic Journal*, *123*(567), 162–194.

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.

Bieber, I. E. (2013). *Frauen in der Politik: Einflussfaktoren auf weibliche Kandidaturen zum Deutschen Bundestag*. Heidelberg: Springer.

BIK–Aschpurwis & Behrens GmbH. (2001). *BIK-Regionen. Ballungsräume, Stadtzentren, Mittel-/Unterzentrengebiete. Methodenbeschreibung zur Aktualisierung [BIK regions. Agglomeration, city centers, medium and small regional centers].* `http://www.bik-gmbh.de/texte/BIKRegionen2000.pdf`, accessed: 2 October 2018.

Bischof, G., & Ambrose, S. E. (1992). *Eisenhower and the German POWs: Facts Against Falsehood*. Baton Rouge: Louisiana State University Press.

Björklund, A., & Jäntti, M. (1997). Intergenerational Income Mobility in Sweden compared to the United States. *The American Economic Review*, *87*(5), 1009–1018.

Bodory, H., Camponovo, L., Huber, M., & Lechner, M. (2018). The Finite Sample Performance of Inference Methods for Propensity Score Matching and Weighting Estimators. *Journal of Business & Economic Statistics*, 1–18.

Brainerd, E. (2017). The Lasting Effect of Sex Ratio Imbalance on Marriage and Family: Evidence from World War II in Russia. *Review of Economics and Statistics*, *99*(2), 229–242.

Bratton, K. A., & Ray, L. P. (2002). Descriptive Representation, Policy Outcomes, and Municipal Day-Care Coverage in Norway. *American Journal of Political Science*, *46*(2), 428-437.

Browning, M., Chiappori, P.-A., & Weiss, Y. (2008). *The Economics of the Family.* `https://www.tau.ac.il/~weiss/fam_econ/`, accessed: 7 November 2018.

Bursztyn, L., Fujiwara, T., & Pallais, A. (2017). 'Acting Wife': Marriage Market Incentives and Labor Market Investments. *American Economic Review*, *107*(11), 3288–3319.

Bursztyn, L., González, A. L., & Yanagizawa-Drott, D. (2018). *Misperceived Social Norms: Female Labor Force Participation in Saudi Arabia.* NBER Working Paper No. 24736.

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

Casas-Arce, P., & Saiz, A. (2015). Women and Power: Unpopular, Unwilling, or Held Back? *Journal of Political Economy*, *123*(3), 641–669.

Chabé-Ferret, B., & Gobbi, P. E. (2018). *Economic Uncertainty and Fertility Cycles: The Case of the Post-WWII Baby Boom.* CEPR Discussion Papers 13374.

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

Chattopadhyay, R., & Duflo, E. (2004). Women as policy makers: Evidence from a randomized policy experiment in India. *Econometrica*, *72*(5), 1409–1443.

Chiappori, P.-A., Fortin, B., & Lacroix, G. (2002). Marriage Market, Divorce Legislation, and Household Labor Supply. *Journal of Political Economy*, *110*(1), 37–72.

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.

Cooke, L. P. (2006). "Doing" gender in context: Household bargaining and risk of divorce in Germany and the United States. *American Journal of Sociology*, *112*(2), 442–472.

Crémieux, P.-Y., Ouellette, P., & Pilon, C. (1999). Health Care Spending as Determinants of Health Outcomes. *Health Economics*, *8*(7), 627–639.

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.

Datenbank VOLKPARL. (n.d.). *Die Abgeordneten der 10. Volkskammer der DDR (VOLKPARL).* `http://zhsf.gesis.org/vk_db/vk_db.php`, accessed: 27 April 2016.

Dee, T. S., & Evans, W. N. (2003). Teen Drinking and Educational Attainment: Evidence

from Two-Sample Instrumental Variables Estimates. *Journal of Labor Economics*, *21*(1), 178–209.

Dell, M., Jones, B. F., & Olken, B. A. (2014). What Do We Learn from the Weather? The New Climate–Economy Literature. *Journal of Economic Literature*, *52*(3), 740–798.

Deloitte. (2018). *Der Deutsche Fitnessmarkt [The German Fitness market]. Benchmarking report.*

Der Bundeswahlleiter. (2016a). *Election to the 11th German Bundestag on 25 January 1987.* `https://www.bundeswahlleiter.de/en/bundestagswahlen/1987.html`, accessed: 23 March 2016.

Der Bundeswahlleiter. (2016b). *Election to the 12th German Bundestag on 2 December 1990.* `https://www.bundeswahlleiter.de/en/bundestagswahlen/1990.html`, accessed: 23 March 2016.

Der Bundeswahlleiter. (2017a). *Election to the 13th German Bundestag on 16 October 1994.* `https://www.bundeswahlleiter.de/en/bundestagswahlen/1994.html`, accessed: 20 December 2017.

Der Bundeswahlleiter. (2017b). *Ergebnisse früherer Bundestagswahlen.* `https://www.bundeswahlleiter.de/en/dam/jcr/397735e3-0585-46f6-a0b5-2c60c5b83de6/btw_ab49_gesamt.pdf`, accessed: 11 June 2018.

Der Bundeswahlleiter. (2018). *Election to the 10th German Bundestag on 6 March 1983.* `https://www.bundeswahlleiter.de/en/bundestagswahlen/1983.html`, accessed: 27 June 2018.

Devereux, P. J., & Hart, R. A. (2010). Forced to be Rich? Returns to Compulsory Schooling in Britain. *The Economic Journal*, *120*(549), 1345–1364.

Doepke, M., & Kindermann, F. (2017). Intrahousehold Decision Making and Fertility. In M. Cervellati & U. Sunde (Eds.), *Demographic Change and Long-run Development.* MIT Press.

Dollar, D., Fisman, R., & Gatti, R. (2001). Are women really the 'fairer' sex? corruption and women in government. *Journal of Economic Behavior & Organization*, *46*(4), 423–429.

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.*

Eagly, A. H., & Karau, S. J. (2002). Role congruity theory of prejudice toward female leaders. *Psychological Review*, *109*(3), 573–598.

Eccles, J. S., & Harold, R. D. (1991). Gender differences in sport involvement: Applying the Eccles' expectancy-value model. *Journal of Applied Sport Psychology*, *3*(1), 7–35.

European Centre for the Development of Vocational Training. (2017). *Germany - 2012 Berlin memorandum on cooperation in VET in Europe brings positive outcome.* `http://www.cedefop.europa.eu/en/news-and-press/news/germany-2012-berlin-memorandum-cooperation-vet-europe-brings-positive-outcome`, accessed: 28 October 2018.

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.

Faust, G. (2006). Zum Stand der Einschulung und der neuen Schuleingangsstufe in Deutschland. *Zeitschrift für Erziehungswissenschaft*, *9*(3), 328–347.

Federal Institute for Population Research. (2018a). *Divorce rates.* `https://www.bib.bund.de/DE/Fakten/Lebensformen/Scheidungsziffern.html`, accessed: 23 April 2019.

Federal Institute for Population Research. (2018b). *Non-marital births in East and West Germany, 1946- 2016.* `https://www.bib.bund.de/DE/Fakten/Fakt/F05-Nichtehelichenquote-West-Ost-ab-1946.html`, accessed: 23 April 2019.

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

Federal Statistical Office. (2016). *Rechnungsergebnisse der öffentlichen Haushalte für*

soziale Sicherung und für Gesundheit, Sport, Erholung – 2010 [Annual financial statement of the public authorities for social security, health, sport, recreation]. Wiesbaden: Statistisches Bundesamt.

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

Federal Statistical Office (Destatis). (2002). *Alle politisch selbständigen Gemeinden mit ausgewählten Merkmalen am 31.12.2001.* https://www.destatis.de/DE/Themen/Laender-Regionen/Regionales/Gemeindeverzeichnis/Administrativ/Archiv/GVAuszugJ/31122001_Auszug_GV.html, accessed: 8 April 2019.

Federal Statistical Office (Destatis). (2003). *Alle politisch selbständigen Gemeinden mit ausgewählten Merkmalen am 31.12.2002.* https://www.destatis.de/DE/Themen/Laender-Regionen/Regionales/Gemeindeverzeichnis/Administrativ/Archiv/GVAuszugJ/31122002_Auszug_GV.html, accessed: 8 April 2019.

Federal Statistical Office (Destatis). (2004). *Alle politisch selbständigen Gemeinden mit ausgewählten Merkmalen am 31.12.2003.* https://www.destatis.de/DE/Themen/Laender-Regionen/Regionales/Gemeindeverzeichnis/Administrativ/Archiv/GVAuszugJ/31122003_Auszug_GV.html, accessed: 8 April 2019.

Federal Statistical Office (Destatis). (2005). *Alle politisch selbständigen Gemeinden mit ausgewählten Merkmalen am 31.12.2004.* https://www.destatis.de/DE/Themen/Laender-Regionen/Regionales/Gemeindeverzeichnis/Administrativ/Archiv/GVAuszugJ/31122004_Auszug_GV.html, accessed: 8 April 2019.

Federal Statistical Office (Destatis). (2006). *Alle politisch selbständigen Gemeinden mit ausgewählten Merkmalen am 31.12.2005.* https://www.destatis.de/DE/Themen/Laender-Regionen/Regionales/Gemeindeverzeichnis/Administrativ/Archiv/GVAuszugJ/31122005_Auszug_GV.html, accessed: 8 April 2019.

Federal Statistical Office (Destatis). (2007). *Alle politisch selbständigen Gemeinden mit ausgewählten Merkmalen am 31.12.2006.* https://www.destatis.de/DE/Themen/

Laender-Regionen/Regionales/Gemeindeverzeichnis/Administrativ/Archiv/GVAuszugJ/31122006_Auszug_GV.html, accessed: 8 April 2019.

Fenn, M., & Körner, H.-M. (2007). Das Schulwesen. In M. Spindler (Ed.), *Handbuch der bayerischen Geschichte Bd. IV,2: Das Neue Bayern: Von 1800 bis zur Gegenwart. Zweiter Teilband: Innere Entwicklung und kulturelles Leben.* Munich: C.H.Beck.

Fernández, R. (2007). Women, Work, and Culture. *Journal of the European Economic Association*, *5*(2-3), 305–332.

Fernández, R. (2013). Cultural Change as Learning: The Evolution of Female Labor Force Participation over a Century. *American Economic Review*, *103*(1), 472–500.

Fernández, R., Fogli, A., & Olivetti, C. (2004). Mothers and Sons: Preference Formation and Female Labor Force Dynamics. *The Quarterly Journal of Economics*, *119*(4), 1249–1299.

Fersterer, J., Pischke, J.-S., & Winter-Ebmer, R. (2008). Returns to Apprenticeship Training in Austria: Evidence from Failed Firms. *Scandinavian Journal of Economics*, *110*(4), 733–753.

Fitzenberger, B., Schnabel, R., & Wunderlich, G. (2004). The gender gap in labor market participation and employment: A cohort analysis for West Germany. *Journal of Population Economics*, *17*(1), 83–116.

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.

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

German Statutory Pension Insurance Scheme. (2017). *Rentenversicherung in Zeitreihen*

[Time series of the German pension insurance scheme]. Berlin.

Gillingham, J. (1986). The "Deproletarianization" of German Society: Vocational Training in the Third Reich. *Journal of Social History*, *19*(3), 423–432.

Göksel, İ. (2013). Female labor force participation in Turkey: The role of conservatism. *Women's Studies International Forum*, *41*, 45–54.

Goldin, C., & Olivetti, C. (2013). Shocking Labor Supply: A Reassessment of the Role of World War II on Women's Labor Supply. *American Economic Review*, *103*(3), 257–62.

Golsteyn, B. H., & Stenberg, A. (2017). Earnings over the Life Course: General versus Vocational Education. *Journal of Human Capital*, *11*(2), 167–212.

Grant, I., Kesternich, I., Steckenleiter, C., & Winter, J. (2018). Historic sex-ratio imbalances predict female participation in the market for politicians. *Journal of Economic Behavior & Organization*, *156*, 144–165.

Grosjean, P. A., & Khattar, R. (2019). It's Raining Men! Hallelujah? The Long-Run Consequences of Male-Biased Sex Ratios. *The Review of Economic Studies*, *86*(2), 723—754.

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

Guinnane, T. W. (2011). The Historical Fertility Transition: A Guide for Economists. *Journal of Economic Literature*, *49*(3), 589–614.

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), PE1077–E1086.

Guttentag, M., & Secord, P. F. (1983). *Too many women? the sex ratio question*. Beverly Hills, CA and London, UK: Sage Publications.

Hahn, J., Todd, P., & Van der Klaauw, W. (2001). Identification and Estimation of Treatment Effects with a Regression-Discontinuity Design. *Econometrica*, *69*(1), 201–209.

Ham, J. C., Swenson, C., İmrohoroğlu, A., & Song, H. (2011). Government programs can improve local labor markets: Evidence from State Enterprise Zones, Federal Em-

powerment Zones and Federal Enterprise Community. *Journal of Public Economics*, *95*(7-8), 779–797.

Handl, J. (1977). Sozio-ökonomischer Status und der Prozess der Statuszuweisung - Entwicklung und Anwendung einer Skala. In J. Handl, K. U. Mayer, & W. Müller (Eds.), *Klassenlagen und Sozialstruktur: empirische Untersuchungen für die Bundesrepublik Deutschland.* Frankfurt, New York: Campus Verlag.

Hanushek, E. A., Schwerdt, G., Woessmann, L., & Zhang, L. (2017). General Education, Vocational Education, and Labor-Market Outcomes over the Lifecycle. *Journal of Human Resources*, *52*(1), 48–87.

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

Harris, A. R., Evans, W. N., & Schwab, R. M. (2001). Education spending in an aging America. *Journal of Public Economics*, *81*(3), 449–472.

Hazan, M., & Maoz, Y. D. (2002). Women's Labor Force Participation and the Dynamics of Tradition. *Economics Letters*, *75*(2), 193–198.

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.

Heineman, E. D. (1999). *What Difference Does a Husband Make? Women and Marital Status in Nazi and Postwar Germany* (Vol. 33). Berkeley and Los Angeles: University of California Press.

Hener, T. (2015). Do Couples Bargain over Fertility? *Annals of Economics and Statistics/Annales d'Économie et de Statistique*(117/118), 211–231.

Henke, K.-D. (1996). *Die amerikanische Besetzung Deutschlands.* Munich: Oldenbourg Verlag.

Hirano, K., & Imbens, G. W. (2004). The propensity score with continuous treatments. In A. Gelman & X.-L. Meng (Eds.), *Applied Bayesian Modeling and Causal Inference from Incomplete-Data Perspectives: An Essential Journey with Donald Rubin's Statistical Family* (pp. 73–84). John Wiley & Sons.

Hitsch, G. J., Hortaçsu, A., & Ariely, D. (2010). Matching and Sorting in Online Dating. *American Economic Review*, *100*(1), 130–63.

Hogan, R. E. (2004). Challenger Emergence, Incumbent Success, and Electoral Accountability in State Legistlative Elections. *The Journal of Politics*, *66*(4), 1283–1303.

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

Hottes, K., & Teubert, R. (1977). *Vertriebene und Flüchtlinge im rheinisch-westfälischen Industriegebiet. Materialien zur Raumordnung aus dem Geographischen Institut der Ruhr-Universität Bochum 18.* Geograph. Inst. d. Ruhr-Univ., Forschungsabt. für Raumordnung.

Huang, H., & Humphreys, B. R. (2012). Sports participation and happiness: Evidence from us microdata. *Journal of Economic Psychology*, *33*(4), 776–793.

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

Huber, M., Lechner, M., & Wunsch, C. (2013). The performance of estimators based on the propensity score. *Journal of Econometrics*, *175*(1), 1–21.

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.

Imbens, G. W., & Lemieux, T. (2008). Regression Discontinuity Designs: A Guide to Practice. *Journal of Econometrics*, *142*(2), 615–635.

infas, DLR. (2010). *Mobilität in Deutschland 2008 – Ergebnisbericht: Struktur, Aufkom-*

men, Emissionen, Trends [Mobility in Germany 2008 – Result report]. Bonn/Berlin: Infas / DLR.

Inoue, A., & Solon, G. (2010). Two-sample Instrumental Variables Estimators. *The Review of Economics and Statistics*, *92*(3), 557–561.

Jackson, C. K., Johnson, R. C., & Persico, C. (2016). The Effects of School Spending on Educational and Economic Outcomes: Evidence from School Finance Reforms. *The Quarterly Journal of Economics*, *131*(1), 157–218.

Jayachandran, S. (2017). Fertility Decline and Missing Women. *American Economic Journal: Applied Economics*, *9*(1), 118–39.

Jürges, H. (2013). Collateral damage: The german food crisis, educational attainment and labor market outcomes of German post-war cohorts. *Journal of Health Economics*, *32*(1), 286–303.

Kaldor, N. (1945). The German War Economy. *The Review of Economic Studies*, *13*(1), 33–52.

Kalmijn, M. (1998). Intermarriage and Homogamy: Causes, Patterns, Trends. *Annual Review of Sociology*, *24*(1), 395–421.

Kesternich, I., Siflinger, B., Smith, J. P., & Steckenleiter, C. (2018). *Unbalanced sex ratios in Germany caused by World War II and their effect on fertility: A life cycle perspective.* CentER Discussion Paper; Vol. 2018-032.

Kesternich, I., Siflinger, B., Smith, J. P., & Winter, J. K. (2014). The Effects of World War II on Economic and Health Outcomes across Europe. *Review of Economics and Statistics*, *96*(1), 103–118.

Kesternich, I., Siflinger, B., Smith, J. P., & Winter, J. K. (2015). Individual Behaviour as a Pathway between Early-life Shocks and Adult Health: Evidence from Hunger Episodes in Post-war Germany. *The Economic Journal*, *125*(588), F372–F393.

Kesternich, I., Smith, J. P., Winter, J. K., & Hörl, M. (in press). Early-Life Circumstances Predict Measures of Trust among Adults: Evidence from Hunger Episodes in Post-War Germany. *The Scandinavian Journal of Economics*.

Kift, D. (2008). Zwischen "eingegliedert werden" und "sich angenommen fühlen". Zur Integration der Flüchtlinge und Vertriebenen in Nordrhein-Westfalen in vergleichender Perspektive. In M. Krauss (Ed.), *Integrationen. Vertriebene in den deutschen*

Ländern nach 1945. Goettingen: Vandenhoeck and Ruprecht.

Kirner, E. (1980). *Ursachen für die Unterschiede in der Höhe der Versichertenrenten an Frauen und an Männer in der gesetzlichen Rentenversicherung* (Vol. 57). Berlin: Duncker & Humblot.

Klevmarken, A. (1982). Missing Variables and Two-Stage Least Squares Estimation from More than One Data Set. *IUI Working Paper, No. 62, The Research Institute of Industrial Economics (IUI), Stockholm.*

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.

Kolinsky, E. (1991). Political participation and parliamentary careers: Women's quotas in West Germany. *West European Politics*, *14*(1), 56–72.

Krauss, M. (1985). "Deutsche sind Deutsche, ... gleichgültig, aus welchem Teil Deutschlands sie stammen." Flüchtlinge und Vertriebene im Trümmermünchen. In F. Prinz & M. Krauss (Eds.), *Trümmerleben. Texte, Dokumente, Bilder aus den Münchner Nachkriegsjahren.* Munich: Deutscher Taschenbuch Verlag.

Kröhnert, S. (2008). *Bevölkerungsentwicklung in Ostdeutschland [Population trends in East Germany].* `http://www.berlin-institut.org/online-handbuchdemografie/bevoelkerungsdynamik/regionale-dynamik/ostdeutschland.html`, Access: September 10, 2017.

Kroener, B. R., Müller, R.-D., & Umbreit, H. (1988). *Organisation und Mobilisierung des deutschen Machtbereichs. Volume 1: Kriegsverwaltung, Wirtschaft und Personelle Ressourcen, 1939-1941.* Stuttgart: Deutsche Verlags-Anstalt.

Ladd, H. F. (1992). Population Growth, Density and the Costs of Providing Public Services. *Urban Studies*, *29*(2), 273–295.

Lafortune, J. (2013). Making yourself attractive: Pre-marital investments and the returns to education in the marriage market. *American Economic Journal: Applied Economics*, *5*(2), 151–178.

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

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

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., Miquel, R., & Wunsch, C. (2011). Long-run effects of public sector sponsored training in west germany. *Journal of the European Economic Association*, *9*(4), 742–784.

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

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

Leitner, S., Ostner, I., & Schmitt, C. (2008). Family policies in Germany. In *Family Policies in the Context of Family Change. The Nordic Countries in Comparative Perspective* (pp. 175–202). Springer.

Lichter, D. T., LeClere, F. B., & McLaughlin, D. K. (1991). Local Marriage Markets and the Marital Behavior of Black and White Women. *American Journal of Sociology*, *96*(4), 843–867.

Lumey, L. H., Stein, A. D., & Susser, E. (2011). Prenatal Famine and Adult Health. *Annual Review of Public Health*, *32*, 237–262.

Malamud, O., & Pop-Eleches, C. (2010). General Education versus Vocational Training: Evidence from an Economy in Transition. *The Review of Economics and Statistics*, *92*(1), 43–60.

Mason, T. W. (1975). *Arbeiterklasse und Volksgemeinschaft. Dokumente und Materialien zur deutschen Arbeiterpolitik 1936 - 1939*. Opladen: Westdeutscher Verlag.

Matland, R. E., & Studlar, D. T. (1996). The Contagion of Women Candidates in Single-

Member District and Proportional Representation Electoral Systems: Canada and Norway. *Journal of Politics*, *58*(3), 707–733.

Mayer, K. U. (1977). Statushierarchie und Heiratsmarkt - Empirische Analysen zur Struktur des Schichungssystem in der Bundesrepublik und zur Ableitung einer Skala des sozialen Status. In J. Handl, K. U. Mayer, & W. Müller (Eds.), *Klassenlagen und Sozialstruktur: empirische Untersuchungen für die Bundesrepublik Deutschland.* Frankfurt, New York: Campus Verlag.

Mayer, K. U. (1979). Berufliche Tätigkeit, berufliche Stellung und beruflicher Status – empirische Vergleiche zum Klassifikationsproblem. In F. U. Pappi (Ed.), *Sozialstrukturanalysen mit Umfragedaten.* Königsstein/Ts: Athenäum Verlag.

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

Meer, J. (2007). Evidence on the Returns to Secondary Vocational Education. *Economics of Education Review*, *26*(5), 559–573.

Minnesota Population Center. (2017). *Integrated Public Use Microdata Series, International: Version 6.5 [dataset] . Minneapolis: University of Minnesota.* `http://doi.org/10.18128/D020.V6.5`, accessed: January 22, 2018.

Moeller, R. G. (1989). Reconstructing the Family in Reconstruction Germany: Women and Social Policy in the Federal Republic, 1949-1955. *Feminist Studies*, *15*(1), 137–169.

Moeller, R. G. (1993). *Protecting motherhood: Women and the Family in the Politics of Postwar West Germany.* Berkeley: University of California Press.

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.

Moscone, F., Knapp, M., & Tosetti, E. (2007). Mental health expenditure in England: A spatial panel approach. *Journal of Health Economics*, *26*(4), 842–864.

Myers, P. S., & Mauldin, P. W. (1952). *Population of the Federal Republic of Germany and West Berlin.* Washington, D.C.: US Government Printing Office.

Neelsen, S., & Stratmann, T. (2011). Effects of prenatal and early life malnutrition: Evidence from the Greek famine. *Journal of Health Economics*, *30*(3), 479–488.

Niederle, M., & Vesterlund, L. (2007). Do Women Shy Away From Competition? Do Men Compete Too Much? *The Quarterly Journal of Economics*, *122*(3), 1067–1101.

Niehuss, M. (1993). Kontinuität und Wandel der Familie in den 50er Jahren. In A. Schildt & A. Sywottek (Eds.), *Modernisierung im Wiederaufbau. Die westdeutsche Gesellschaft der 50er Jahre.* Bonn: Forschungsinstitut der Friedrich Ebert-Stiftung.

Niehuss, M. (2001). *Familie, Frau und Gesellschaft. Studien zur Strukturgeschichte der Familie in Westdeutschland 1945–1946.* Göttingen: Vandenhoeck and Ruprecht.

Oosterbeek, H., & Webbink, D. (2007). Wage Effects of an extra year of basic vocational education. *Economics of Education Review*, *26*(4), 408–419.

Oreffice, S., & Quintana-Domeque, C. (2010). Anthropometry and socioeconomics among couples: Evidence in the United States. *Economics & Human Biology*, *8*(3), 373–384.

Overmans, R. (2004). *Deutsche militärische Verluste im Zweiten Weltkrieg.* Munich: Oldenbourg Verlag.

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)

Pfau-Effinger, B. (2004). Socio-historical paths of the male breadwinner model–an explanation of cross-national differences. *The British Journal of Sociology*, *55*(3), 377–399.

Powley, E. (2006). *Rwanda: The Impact of Women Legislators on Policy Outcomes Affecting Children and Families: The State of the World's Children 2007 Background Paper.* United nations children's fund (UNICEF). Division of policy and planning.

Rass, C., & Rohrkamp, R. (2007). *Deutsche Soldaten 1939-1945: Handbuch einer biographischen Datenbank zu Mannschaften und Unteroffizieren von Heer, Luftwaffe und Waffen-SS.* `http://darwin.bth.rwth-aachen.de/opus3/volltexte/2007/`

2120/pdf/2120.pdf/, accessed: 27 June 2017.

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.

Richter, D., & Schupp, J. (2015). The SOEP Innovation Sample (SOEP IS). *Schmollers Jahrbuch: Journal of Applied Social Science Studies/Zeitschrift für Wirtschafts-und Sozialwissenschaften*, *135*(3), 389–400.

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 Subclassification 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* (Unpublished master's thesis). Mathematical Statistics, Stockholm University.

Schimpl-Neimanns, B., & Frenzel, H. (1995). *1-Prozent-Stichprobe der Volks-und Berufszählung 1970: Datei mit Haushalts- und Familiennummern und revidierter Teilstichprobe für West-Berlin. Dokumentation und Datenaufbereitung.* (ZUMA-

Technischer Berich 95/06. Mannheim)

Schulz, G. (1998). Soziale Sicherung von Frauen und Familien. In H. G. Hockerts (Ed.), *Drei Wege deutscher Sozialstaatlichkeit. NS-Diktatur, Bundesrepublik und DDR im Vergleich.* München: R. Oldenbourg Verlag.

Schüttoff, U., Pawlowski, T., Downward, P., & Lechner, M. (2018). Sports Participation and Social Capital Formation During Adolescence. *Social Science Quarterly*, *99*(2), 683–698.

Scientific Research Service of the German Bundestag. (2011). *Deutsche Kriegsgefangene in sowjetischem Gewahrsam [German prisoners of war in war imprisonment in the Soviet Union].* `https://www.bundestag.de/resource/blob/413294/.../WD-1-011-11-pdf-data.pdf`, accessed April 16, 2019.

Seebens, H. (2006). Bargaining over Fertility in Rural Ethiopia. *Proceedings of the German Development Economics Conference, Berlin 2006 / Verein für Socialpolitik, Research Committee Development Economics, No. 25*.

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

Statistisches Bundesamt. (1950). Erstes Ergebnis der Volkszählung 1950. Vorläufige Bevölkerungszahlen der Länder des Bundesgebiets. *Wirtschaft und Statistik*, *2*(7), 246–248.

Statistisches Bundesamt. (1955). II: Bevölkerungsbewegung, 1. Eheschliessungen, Geborene und Gestorbene 1871 bis 1953. In *Statistisches Jahrbuch für die Bundesrepublik Deutschland 1954.* Stuttgart-Köln: W. Kohlhammer.

Statistisches Bundesamt. (1967). Die unehelichen Kinder. *Wirtschaft und Statistik*, *19*(7), 391-394.

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

Statistisches Reichsamt. (1937). IX. Löhne und Gehälter. a. Tatsächliche Arbeitsverdienste. 3. Durchschnittsverdienste in 16 Gewerben im Jahre 1936 nach den Ergebnissen der amtlichen Lohnerhebungen. In *Statistisches Jahrbuch für das Deutsche Reich 1937.* Berlin: Verlag für Sozialpolitik, Wirtschaft und Statistik, Paul Schmidt.

Statistisches Reichsamt. (1940). Weitere Teilergebnisse der Volks- und Berufszählung vom 17. Mai 1939. Die ständige Bevölkerung nach Geschlecht, Alter und Familienstand. *Wirtschaft und Statistik*, *20*(15).

Statistisches Reichsamt. (1941). *Volkszählung. Die Bevölkerung des Deutschen Reichs nach den Ergebnissen der Volkszählung 1939. Heft 2: Geschlecht, Alter und Familienstand der Bevölkerung des Deutschen Reichs. Tabellenteil.* Berlin: Verlag für Sozialpolitik, Wirtschaft und Statistik, Paul Schmidt.

Steckenleiter, C., Lechner, M., Pawlowski, T., & Schütoff, U. (2019). *Do local public expenditures on sports facilities affect sports participation in Germany?* (mimeo)

Steinberg, H. G. (1991). *Die Bevölkerungsentwicklung in Deutschland im Zweiten Weltkrieg mit einem Überblick über die Entwicklung von 1945 bis 1990.* Bonn: Kulturstiftung der deutschen Vertriebenen.

Steinhauer, A. (2018). *Working Moms, Childlessness, and Female Identity.* Sciences Po LIEPP Working Paper Nr.79.

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.

Stevenson, B. (2010). Beyond the Classroom: Using Title IX to Measure the Return to High School Sports . *The Review of Economics and Statistics*, *92*(2), 284–301.

Tegtmeyer, H. (1979). Berufliche und soziale Umschichtung der Bevölkerung: Methodische Anmerkungen zur Planung, Durchführung und Aufbereitung der Befragung. In H. Tegtmeyer (Ed.), *Soziale Strukturen und individuelle Mobilität. Beiträge zur sozio-demographischen Analyse der Bundesrepublik Deutschland. Schriftenreihe des Bundesinstituts für Bevölkerungsforschung.* Wiesbaden: VS Verlag für Sozialwissenschaften.

Teso, E. (2019). The Long-Term Effect of Demographic Shocks on the Evolution of Gender Roles: Evidence from the Transatlantic Slave Trade. *Journal of the European Economic Association*, *17*, 497—534.

Tiebout, C. M. (1956). A Pure Theory of Local Expenditures. *Journal of Political Economy*, *64*(5), 416–424.

TNS Infratest. (2017). *Alterssicherung in Deutschland 2015 (ASID 2015). Tabellen-*

band 3 – Deutschland , München. Erschienen als Forschungsbericht des BMAS 474/D. https://www.bmas.de/SharedDocs/Downloads/DE/PDF-Publikationen/Forschungsberichte/fb-474d-alterssicherung-deutschland-2015.pdf?__blob=publicationFile&v=2, accessed: 28 October 2018.

United Nations Women. (n.d.). *Facts and Figures: Leadership and Political Participation.* http://www.unwomen.org/en/what-we-do/leadership-and-political-participation/facts-and-figures, accessed: January 4, 2017.

Varanasi, N. (2009). *Intrahousehold bargaining and fertility.* (Unpublished manuscript)

Verfassungen. (2004). *Kontrollratsgesetz Nr. 18 Wohnungsgesetz.* http://www.verfassungen.de/de/de45-49/kr-gesetz18.htm, accessed: August 7, 2017.

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.

Wassermann, L. (2004). All of Statistics : A Concise Course in Statistical Inference. *Statistical inference. Springer, Berlin.*

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.

Wolsing, T. (1977). *Untersuchungen zur Berufsausbildung im Dritten Reich* (K. Ottenheym & H. Schallenberger, Eds.). Kastellaun: Aloys Henn Verlag.

World Bank. (2018). *World Development Indicators 2018.* The World Bank.

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

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

Zimmerman, K. (1993). Labour Responses to Taxes and Benefits in Germany. In A. B. Atkinson & G. V. Mogensen (Eds.), *Welfare and Work Incentives: A Northern European Perspective* (pp. 192–240). Oxford University Press.

Curriculum Vitae

Education

2015 – 2020	Ph.D. in Economics and Finance, University of St. Gallen Thesis supervisor: Prof. Dr. Michael Lechner
2012 – 2014	M.Sc. in Economics, Ludwig-Maximilians-Universität München
2013 – 2014	Visiting Student, M.Sc. in Economics, Université de Lausanne
2012	American University, Washington Semester Program, Global Economics & Business, Washington, D.C.
2009 – 2012	B.Sc. in Economics, Ludwig-Maximilians-Universität München

Professional experience

2014 – 2019	Research assistant, Chair Prof. Dr. Michael Lechner, Swiss Institute for Empirical Economic Research, University of St. Gallen
2014	Student research assistant, Chair Prof. Dr. Monika Schnitzer, Seminar for Comparative Economics, Ludwig-Maximilians-Universität München