Long-run labour market effects of individual sports activities

Michael Lechner

April 2008 Discussion Paper no. 2008-13

Long-run labour market effects of

individual sports activities

Michael Lechner

Author's address: Prof. Dr. Michael Lechner
                          SEW-HSG
                          Varnbüelstrasse 14
                          9000 St. Gallen
                          Phone   +41 71 224 28 14
                          Fax   +41 71 224 23 02
                          Email   Michael.lechner@unisg.ch
                          Website   www.sew.unisg.ch

Abstract

This microeconometric study analyzes the effects of individual leisure sports participation on long-term labour market variables, on socio-demographic as well as on health and subjective well-being indicators for West Germany based on individual data from the German Socio-Economic Panel study (GSOEP) 1984 to 2006. Econometric problems due to individuals choosing their own level of sports activities are tackled by combining informative data and flexible semiparametric estimation methods with a specific way to use the panel dimension of the data. The paper shows that sports activities have sizeable positive long-term labour market effects in terms of earnings and wages, as well as positive effects on health and subjective well-being.

Keywords

1 Introduction

The positive effect of physical activities on individual health is widely acknowledged both in academics and the general public. Nevertheless, there is still a substantial part of the population that is not actively involved in sports. For example, in Germany about 40% of the population older than 18 does not participate in sports activities at all, which is about the average for Europe (they tend to be lower in Southern and higher in Northern Europe). A similar pattern appears in the USA.[1] These non-activity figures are surprisingly high considering that many Western countries subsidize the leisure sports sector substantially.[2] The large subsidies are justified by considerable positive externalities participation in sports may have, for example by increasing public health and fostering social integration of migrants or other social groups, who otherwise deal with integration difficulties (for Germany, see Deutscher Bundestag, 2006; for Austria, see Weiss and Hilscher, 2003).

In this paper, the focus is on the effects of individual participation in leisure time sports on individual labour market outcomes in the long run. Intuitively, one might expect that such labour market effects usually result from one or several of the following three channels. The first channel relates to direct productivity effects. Improved health and improved individual well-being might lead to direct gains in individual productivity that is rewarded in the la-

---

[1] The figures for Germany are taken from Bundestag (2006, p. 94). The source for the European numbers is Gratton and Taylor, (2000, chapter 5), while the US figure comes from Ruhm (2000) and Wellman and Friedberg (2002). The US figures are based on a broader definition of activities than the European ones including general physical activities. According to that definition, about 25-30% of the relevant adult US population does not engage in leisure physical activities including sports.

[2] Public expenditures come in various forms and from various levels of government. They may be directed to investments in infrastructure and the subsidisation of sports organisations, information campaigns, tax rebates for sports related expenditures (in particular donations), etc. The relative importance of the different expenditure categories and the overall amounts, as well as the way how the support system is organized varies drastically from one country to another (see Gratton and Taylor, 2000). In addition, health organisations and firms invest in encouraging people to take up physical activities. This diversity of sponsoring institutions and types of expenditures makes it extremely difficult to get a reliable estimate of the total expenditures for non-professional sports.

bour market. The second channel is made up of social networking effects that are particularly relevant for sport activities performed in groups. As for a third channel sport activities might signal to potential employers that individuals enjoy good health, are motivated and thus will perform well on the job. The paper clearly concentrates on the first channel, although it will be difficult in the empirical analysis to clearly differentiate between the different explanations for the effects found.

To be more precise, this paper addresses two issues that are important to both the individual as well as the public: The first question is whether the health gains appearing in medical studies are still observable when taking a long-run perspective. It is conceivable that the health gains disappear, because the additional 'health capital' may be 'invested' in less healthy activities such as working harder on the job. This of course would put into question on one of the main justifications for the public subsidies. Second, even if the direct health effects are absent in the long run, participation in sports may increase individual productivity which appears desirable as well. Such an increase would be observable in standard labour market outcomes like earnings, wages, and labour supply. Actually identifying such effects would be valuable information that could be used in public information campaigns to increase participation in leisure sports.

There are at least four strands of the literature relevant for this topic. The first strand appears in labour economics and analyzes the effects of participating in high school sports on future labour market outcomes. Based on various data sets mainly from the USA and various econometric methods to overcome the problem of self-selection into high school sports, this literature broadly agrees that participation in such type of sports improves future labour market outcomes (e.g., Barron, Ewing, Waddell, 2000, Ewing, 1998, 2007, Long and Caudill,

2001, Persico, Postlewaite, and Silverman, 2004, and Stevenson 2006, for the USA, and Cornelissen and Pfeifer, 2007, for Germany).[3]

Next, the positive effect of sports activity on physical health is well documented in the medical and epidemiological literature (e.g., Hollmann, Rost, Liesen, Dufaux, Heck, Mader, 1981, Lüschen, Abel, Cockerham, and Kunz, 1993, US Department of Health and Human Services, 1996, Weiss and Hilscher, 2003). There is recent microeconometric evidence of a positive relationship as well: Rashad (2007) analyzes the effects of cycling on health outcomes. Lakdawalla and Philipson (2007) find that physical activity at work reduces body weight and thus the probability of obesity. Bleich, Cutler, Murray, and Adams (2007) look at the relationship of physical activity and the problem of obesity as well. However, they find that the international trend of increasing obesity is more related to changes in how and what people eat than to reductions in physical activity. This finding is somewhat in contrast to previous findings in the medical literature suggesting a more important role of declining physical activity over time (e.g., Prentice and Jebb, 1995). Recent papers, for example Gomez-Pinilla (2008), also suggest that sports activities have a considerable positive effect on mental health.

In addition, there exists a literature linking health and labour market outcomes: Declining health reduces productivity and as a consequence it reduces wages and might reduce labour market participation. An important channel is the impact of body weight, in particular obesity, on labour market outcomes. Obesity is becoming wide spread (e.g., Andreyeva, Michaud, and van Soest, 2005). It increases the risk of mortality, diabetes, high

---

[3] For a related analysis of the effect of high school sports participation on suicides, see Sabo, Miller, Melnick, Farrell, and Barnes (2005).

blood pressure, asthma, and other diseases, and thus drastically reduces labour productivity (e.g., Wellman, and Friedberg, 2002, and the many references given in Ruhm, 2007).

From a policy perspective, it is stressed (e.g., Deutscher Bundestag, 2006) that an important channel of how participation in sports, particularly team sports, may improve future labour market performance is by increasing social skills. Therefore, the sociological literature describing how social capital may improve labour market performance (e.g., Aguilera and Barnabé, 2005) and how 'positive' extracurricular activities in youth lead to more successful labour market performance in later years (e.g., Eccles, Barber, Stone, and Hunt, 2003) is relevant as well.

Despite the large literature on the topics mentioned above, as of yet there appears to be no information available on the effects of leisure sports on individual labour market outcomes. In that the effects of sports on labour market success take time to materialise, estimating long-run effects is particularly relevant. Uncovering such long-run effects, however, comes with particular challenges: The first challenge is the data, which should record individual information over a sufficiently long time. This data should contain measurements of sports activities, labour market success and other outcome variables of interest, as well as the variables that jointly influence the outcomes of interest as well as the decision about participating in sports. In Section 2 and 3, it is argued that the German Socio-Economic Panel Study (GSOEP) with annual measurements from 1984 to currently 2006 could be used for such an analysis, although it suffers from some drawbacks as well.

The second challenge concerns the problem of individual self-selection into different levels of sports activity. For example, if those individuals on well-paying jobs choose higher levels of sports activity, then a comparison of the labour market outcomes of individuals with low and high sports activity levels will not only contain the effects of different activity levels, but may also reflect differences of these groups with regard to other dimensions. This is called

3

the problem of 'selection bias' in the econometric literature (see Heckman, LaLonde, and Smith, 1999), and 'confounding' in the statistical literature (e.g. Rubin, 1974). The fact that selection into sports is not random is well documented, for example, by Becker, Klein, and Schneider (2006) and Schneider and Becker (2005) for Germany, and by Farrell and Shields (2002) for England. However, solving this problem in the usual way, which means conditioning on the variables that pick up these confounding differences may not solve the problem as the values of these conditioning variables may depend on past participation in sports (endogeneity problem of control variables).

In this paper, this endogeneity problem is solved by using a flexible semiparametric econometric estimation technique (a specific variant of a so-called matching estimator) together with performing the analysis in subsamples defined such that in each subsample all individuals have the same level of past sports activity. Then, within each subsample the effects of the next subsequent change in these levels are analyzed. This approach removes (most of) the endogeneity problem as the control (confounding) variables are measured in a period when everybody has the same level of sports activity and their measurement can therefore not be influenced by differences in activities.

The paper intents to contribute to the literature in three dimensions: The first goal is to learn more about the correlates of sports activities by using the GSOEP data with its wealth of information. Since this is done in such a way that the problem of endogeneity is eliminated or at least reduced, the interpretation of the results should be less controversial than in previous studies. The second and main contribution of this study is to uncover the long-run effects of participation in sports on labour market success and several other socio-demographic and health variables. Finally, a methodological point is made by adapting existing semiparametric econometric estimation methods to the specific panel data situation without having to impose

the restrictive assumptions that the popular fixed and random effects panel data estimators would imply.

The results of the analysis of the leisure sports activities selection process suggest that participation in sports is higher for men than for women. They are much lower for non-Germans, particularly for non-German women. Sports activities increase with education, earnings, and 'job quality'. Marriage and children (for women) as well as an older age are associated with a lower involvement in sports.

The analysis of the effects of sports activities on outcomes revealed sizeable labour market effects. As a rough estimate, active participation in sports increases earning by about 1.200 EUR p.a. over a 16 year period compared to no or very low participation in sports. The results translate to rates of return of sports activities in a range of 5% to 10%, suggesting similar magnitudes than for one additional year of schooling. Increased health and improved well-being in general seem to be relevant channels to foster these earnings gains.

The next section analyzes the correlates of the participation in sports activities. It describes the data and the endogeneity problem. Section 3 describes the econometric approach to identify and estimate the effects of sports on the various outcome variables taken into consideration. Section 4 contains the main results and checks of robustness. Section 5 concludes. Appendix A discusses a couple of data related issues. Appendix B describes the procedures used for estimation and inference. For the sake of brevity, additional have been set aside in a second appendix that is available in the internet (www.sew.unisg.ch/lechner/sports_GSOEP).

2   Who participates in leisure sports activities?

2.1   Previous results

As mentioned above, there seems to be common agreement in the literature that sports activities tend to decrease with age, tend to increase with earnings or social status, and that men are more active than women. However, although not much is known in general on further determinants of participation in sports, there are some studies based on individual data that at least give some hints to further factors.

Based on the British Health and Lifestyle Survey with interviews around 1984, Gratton and Taylor (2000) use a logit analysis for sports participation. They report in addition negative associations for past illnesses. Furthermore, they find positive associations of sports participation and not working full-time, as well as for sports participation and being separated or divorced. In a more recent study based on the Health Survey for England conducted in 1997, Farrel and Shields (2002) roughly confirm these findings using a probit model for sports participation. They further point to a negative association of sports participation and the presence of young children, as well as to a positive association related to the presence of older children for men. Furthermore, being a drinker, being white, and not being a smoker is also positively associated with sports participation.

Schneider and Becker (2005) use a binary logit model and the German National Health survey with interviews between 1997 and 1999 for a similar analysis. They confirm the previous findings, except with respect to smoking. They further find that being more satisfied with life in general, having a lower body mass index (BMI), and having received medical advice on physical activity is also positively associated with sports participation. In similar work, Becker, Klein, and Schneider (2006) analyze the 2003 cross-section of the GSOEP. In addition to the 'usual' findings concerning education and age, they find that for

2003 women are more likely than men, and never-married singles are more likely than people who are or have been married to participate in sports. They also find a negative correlation for being a foreigner. Furthermore, they detect correlations for some subjective variables on social networks, subjective and objective health variables, as well as variables capturing policy interest, and general life satisfaction (all measured simultaneously with sports participation) that are correlated.

However, how to interpret the results of these cross-sectional studies is not obvious because they relate a phenomenon (sports activity) that could have been going on for a long time to other variables that may be influenced by past and present sports activities as well. For example, in the study by Becker, Klein, and Schneider (2006) it is not at all clear whether good health increases sports activity or sports activity improves health. The same problem holds for some of the other time varying variables. This gives raise to the so-called endogeneity or reverse causality problem which makes a causal interpretation of the correlates identified in such studies difficult. In the following section, we suggest to use panel data to considerably reduce, if not eliminate, this problem.

2.2 The endogeneity problem reconsidered when panel data are available

In a cross-sectional study, the different sports participation statuses of the individuals have to be related to covariates measured at the same time as the participation status. Therefore, the measurement of the time varying variables in a particular period may already be influenced by current or past sports participation. If we were able to observe values of those variables as they were realized for a specific sports participation status, such values would not be subject to the endogeneity problem as they are not influenced by the actual realisation of the sports participation (i.e. the values of past labour market experience had the individual not participated in sports activities). However, as for every individual we observe only the values of the covariates along with specific realized sports participation. Such (partly counterfactual)

values are not available in a cross-section. This is particularly so, in that the variation in the sports participation status is needed to be able to analyze its determinants.

With panel data it is possible to circumvent this problem by exploiting both the variation of the sports status over time as well as over individuals. 'Determinants' of sports status should be measured close, but prior, to the sports participation decision (as future events do not influence past events). Therefore, the endogeneity problem is resolved, if the analysis is based on individuals who are in the same sports status in the period before the specific sports participation decision is analyzed, and measurements of the covariates prior to that period are available. Thus, using some standard cross-sectional binary choice model for such a specific subsample with the sports participation status of the current period as the dependent variable and last periods' measurements of the covariates as independent variables, leads to considerably more credible results than those obtained from a cross-section.[4] Of course, the drawback is that the conclusions are valid only for the specific population with the particular sports participation status. However, this can be resolved by considering all such populations one-by-one (and taking appropriate averages if desired).

2.3   Findings based on the German Socio-Economic Panel

2.3.1   The data

The German Socio-Economic Panel Study (GSOEP) is a representative panel study with annual measurements starting in 1984. This study uses data from 1984 to 2006. The

---

[4]   In the econometric implementation, I refrain from using off-the-shelf panel econometric models, i.e. in this case fixed effects or random effects models, because they require a considerable number of undesirable assumptions, like strict exogeneity of the regressors and rely more importantly on functional form assumptions for identification that restrain the effects of heterogeneity and imply other important underlying behavioural restrictions. Those restrictions become particularly pronounced for nonlinear models, like logit or probit, which may be required by the nature of the outcome variable that renders a linear specification unattractive. See Lechner, Magnac, and Lollivier (2008) for an overview of the classical nonlinear models for panel data.

GSOEP is interviewer based and recently switched to computer assisted personal interviews (CAPI). It started in West Germany. In 1990 it began including East Germany as well. The GSOEP is one of the work-horses of socio-economic research in Germany, and beyond. More details on the survey and its development can be found in Wagner, Frick, and Schupp (2007) and on the GSOEP website (www.diw.de/gsoep). Details about key questions used in the empirical analysis can be found in Appendix A.1.

Since it is the goal of the empirical analysis to investigate the long-run labour market effects of participation in sports, it is required that in the year of the decision individuals should be aged between 18 and 45. The upper age limit is defined such that there is a considerable chance that individuals are still working at the end of the observation period for the outcomes which last 16 years.[5] Again, in order to measure long-run outcomes as well as pre-decision control variables, the focus is on the West German subsample and on sports participation decisions in the years 1985, 1986, 1988, and 1990 only.[6] All variables are then redefined relative to the respective year of the decision (e.g., for a decision in 1990, the out-come '16 years later' would be taken from the 2006 survey, whereas the 'control' variables, including previous sports activity levels, would in most cases be taken from the 1989 survey). Investigating those four decision periods separately (conditional on the previous sports participation status) would lead to very imprecise estimate due to the small subsample sizes. Therefore, using the redefined variables, the four different starting cohorts are pooled. In

---

[5] Increasing the lower age limit to 24 years leads to similar results, but there is a loss of precision due to the smaller sample size. Defining 16 years as the desired window for measuring long-run effects is of course arbitrary and may be seen as a lower bound for the real long-run effects. There is a trade-off between sample size and the length of the observation window. Since the 2006 survey is the last one available, using 16 years allows analyzing sports activities until 1990. Increasing the observation period further would require using decisions prior to 1990 only and thus reducing sample size further. Since section 4 will show that the precision of the estimates is already an issue, it appears that any further reduction of the sample size comes at a high price too high for the additional gain of up to five more years.

[6] For the West, the years 1987 and 1989 are omitted due to data limitations regarding the sports variable.

other words, if the individuals have the same the same prior sports participation status (and gender) they are pooled irrespective of in which of the four periods they originate. Furthermore, to be consistent with the sections discussing the empirical estimates of the effects of sport, only the results of a balanced panel are reported.[7] Moreover, individuals indicating that they were hospitalized either in the year of the decision or in the year before are not taken into consideration to avoid basing results on seriously ill people, who are expected to participate in sports for other reasons, if at all. As an unavoidable side effect, this rule excludes most women giving birth in those two years. See Appendix A.2 for more details on the sample selection rules.

Participation in sports is measured in four different categories (at least every week, at least every month but not every week, less often than every month, none; see Appendix A.1 for the specific questions in the survey). Table 2.1 shows the development of that variable over time for the combined sample (not yet rearranged relative to the decision years) to get an idea about the dynamics of sports participations in general.

In 1985 35% of the men and 50% of the women did not participate in any sports, whereas 36% of the men and 26% of the women were active on a weekly basis. However, in 2005, these gender differences disappeared: Although slightly more women than men did not participate in any activity (40% compared to 37%), fewer men than women (32% compared to 37%) are active at least on a weekly basis. Thus, while the women in the sample increased their activity levels, the activity levels for men remained fairly constant over time. Becker,

---

[7] To be precise, it is required to be observed in the years -1, 0, 1 to 16 (0 denotes the year of the participation decision, -1 the year before, etc.). The results for a corresponding unbalanced panel requiring only to be observed in the years -1 and 0 are available on request. They support the findings presented in this paper. Using the 'observability' of an individual up to 16 years after the sports participation decision analysed as an outcome variable when evaluating the effects of sports activities does not reveal any effect of activity levels on observability, indicating that the analysis can be conducted on the balanced sample without having to worry to much about attrition bias.

Klein, and Schneider (2006) find similar trends using GSOEP data starting 1992. However, the activity levels they observe are lower, because they base their analysis on a broader definition of the underlying population. It is also important to note that in some years the sports question is based on a five point scale instead of the four point scale. In those years, it appears that people avoid the 'extremes' of the scale more frequently. This pattern has also been observed by Breuer (2004).

Table 2.1: Trends of sports participation over time for men and women (balanced sample)

	Men				Women			
	Frequency of leisure sports activities							
	weekly	monthly	< monthly	none	weekly	monthly	< monthly	none
1985	36	8	21	35	26	6	18	50
1986	38	7	19	35	27	6	17	50
1988	36	8	19	37	27	6	18	49
1990	38	11	26	25	32	9	23	36
1992	32	11	22	36	27	6	20	47
1994	31	9	23	36	26	7	20	47
1995	36	9	24	31	32	8	22	38
1996	32	9	24	35	27	7	21	44
1997	31	9	23	38	28	6	19	46
1998	33	11	25	31	32	7	24	37
1999	29	10	23	37	29	7	18	47
2001	30	9	21	40	32	5	17	46
2003	33	10	27	30	41	5	18	36
2005	32	9	21	37	36	6	18	40

Note: In 1990, 1995, 1998, and 2003 a five point scale is used which splits the category *weekly* into *weekly* and *daily*. For those years the entries in the columns headed by *weekly* include the additional category *daily*.

The empirical analysis will aggregate the four (to five) groups of information on sports activity into two groups only for two reasons: (i) the subsamples within the four (to five) groups are too small for any robust (semiparametric) econometric analysis, which means that the lack of observation would require the reliance on functional form assumptions relating (and restricting) the different effects for the subgroups instead. In this paper, I want to explicitly avoid such restrictions and their undesirable impact on the results (see the discussion in Section 3). (ii) When the five point scale is used instead of the four point scale, different categories appear as extreme categories. The aggregation of all extreme categories into neighbouring categories should be very helpful to mitigate these problems. Thus, following

the medical literature on analysing sports participation from GSOEP data (e.g., Becker, Klein, and Schneider, 2006), from now on, we differentiate between only two levels of activity, namely being active at least monthly and being active less than monthly.

Based on this definition of sports activity, the empirical analysis uses two subsamples of the West German population. The *no-sports sample* consists of those individuals who did not participate in sports at least monthly in the year before the decision is analyzed (year '-1'). The *sports sample* is made up of all individuals reporting at least monthly involvement in sports activities.[8] Furthermore, since the literature suggests substantial differences between men and women, the empirical analysis is stratified by sex.

Using these definitions and sample restrictions, in the *no-sports sample* there are 2027 men and 2338 women, of whom 482 men and 448 women increased their sports activities in the next period above the threshold. In the *sports sample*, out of the 1471 men and 915 women, 339 men and 262 women reduced their sports activities in the next period below the threshold. It is already apparent from these numbers that in the period from 1985 to 1990, men are more likely to participate in sports than women.

2.3.2   Results

Table 2.2 presents sample means of the various covariates for the four different samples stratified according to the sports status in the year analyzed and sex. Thus, pair-wise comparisons of columns (2) vs. (3), (5) vs. (6), (8) vs. (9), and (11) vs. (12) allows to assess the covariate differences that come with the different sports participation status within each subsample. An additional measure to assess the relevance of specific covariates are the coefficients of a binary probit model with sports participation as dependent variable that are pre-

---

[8]   To assess the sensitivity of these decisions, they have been varied to assess the sensitivity of the results with respect on how to define sports participation (see Section 4.3).

sented in columns (4), (7), (10), and (13). To avoid flooding the reader with numbers, coefficients not significant at least at the 10% level are not listed (empty cell in table). When specific variables are omitted from the probit specification, it is usually because either they have been chosen as being part of the reference category (denoted by 'R'), the cell counts are too small, or they do not play a role in the specific subpopulation ('-').[9] Note that comparing columns (2), (3), (5), and (6) of the no-sports sample to the corresponding columns (8), (9), (11), and (12) of the sports sample also gives an indication as to variables *correlated* with sport-participation.[10]

Next, the different groups of variables are considered in turn. First, the cohort dummies capturing the year of the pooled participation decisions indicate that participation in sports is increasing over time. This finding is consistent with results of Table 2.1, as well as with the literature mentioned above.

The next block of variables is related to the socio-demographic situation. The results show that for the no-sports sample, younger individuals are more likely to be active. The coefficients of the probit indicate, however, that this relation is probably non-linear. No such relation appears in the sports sample. The relationship between sports activity and nationality is clear-cut for women: Non-Germans are less likely to be observed as active participants in sports (confirming the findings by Becker, Klein, and Schneider, 2006). For men, this relation seems to exist as well, but is less pronounced, particularly for men in the no-sports sample. In addition, being married is associated with lower sports activity in the no-sports sample. The

---

[9] To support these probit specifications, tests for omitted variables, as well as further general specifications tests against non-normality and heteroscedasticity are conducted. These respective test statistics do not point to serious violations of the statistical assumptions underlying the probit model. They are available on request from the author.

[10] As the sport status used to define the subsamples and the control variables are measured at the same time, such a comparison is only informative about the correlation of sports participation with covariates, not about any causal connection.

effects for the sports sample are smaller for men and absent for women, thus moderating the findings by Becker, Klein, and Schneider (2006). A relationship between divorce and sports activities as reported by Gratton and Taylor (2000) appears to be absent. Finally, the existence of young children in the household of is generally related to a lower level of sports activities of women (as in Farrel and Shields, 2002), whereas the effects for men appear to be absent.[11]

The educational information, which is known from other studies to play an important role, is described by several variables related to formal schooling as well as to vocational education. The results of Table 2.2 support the general finding that sports activities increase with education. This is also in line with a positive association of individual and family earnings with sports participation for women. This relation seems to be almost absent for men casting some doubt on the findings of the literature so far. The same pattern appears for the crude wealth indicator that could be used for this analysis, namely whether the current apartment or house is owned or rented. For men, there are no, or only small differences, whereas more wealthy women are more likely to participate in sports activities.

For those who worked in the year before they started their sports participation, various variables in addition to earnings are also included to characterize the firm (size, sector), the job (duration, earnings, hours, required vocational education, sector, type of occupation, prestige of occupation measured by the Treimann scale, 'autonomy' of occupation measured by a 5 point scale, job position).[12] For those individuals not working, their current status is known as well (unemployed, out of labour force, retiree, students, etc.). Furthermore, there is information on job histories, such as total duration in full-time or part-time employment, and so on.

---

[11] Further socio-demographic information, such as immigration information, etc., has been considered in the estimation but not presented in the table, because they have no further explanatory power in the probit (conditional on the variables already included).

[12] Only selected variables appear in the table.

The results for these durations are however difficult to interpret as they are by definition positively correlated with age.

Table 2.2: Selected descriptive statistics and probit coefficients for the selection process into sports activities

	Sports activity before											
	Less than monthly						At least monthly					
	Men			Women			Men			Women		
	Mean in subsample		Pro-bit	Mean in subsample		Pro-bit	Mean in subsample		Pro-bit	Mean in subsample		Pro-bit
Characteristics	Sport	No S.	S-NS	Sport	No S.	S-NS	Sport	No S.	S-NS	Sport	No S.	S-NS
(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)
Year of sports participation considered (in %)												
1985	36	30	R	33	31	R	21	15	R	18	14	R
1986	21	27	-.21*	20	27		27	30	-.81**	29	27	-.60**
1988	20	23		22	23		25	32	-.90**	25	34	-.84**
1990	23	20		25	19	.29**	27	23	-.65**	28	25	-.55**
Socio-demographic characteristics												
Age in years	31	33		31	33	-.03**	30	31		32	31	
Age: 18-25 (dummy)	.29	.21	.33*	.28	.22	.25*	.31	.31		.27	.29	
German nationality	.76	.75		.91	.69	.51**	.85	.75		.98	.90	.82**
Married	.57	.65	-.52*	.58	.72		.47	.52		.56	.56	
Divorced	.03	.03		.06	.05		.04	.04		.05	.06	
# of kids in household	1.0	1.3		.86	1.2		.76	.85		.83	.82	
Mother of kids age < 3	-	-	-	.13	.18	-.20	-	-	-	.08	.17	-.65**
Mother of kids age < 7	-	-	-	.40	.48	.23*	-	-	-	.33	.38	
Mother of kids age < 10	-	-	-	.54	.70	-.17**	-	-	-	.51	.53	.26*
Education (in %)												
Lower secondary school or no degree	45	50	R	42	57	R	39	42	R	56	61	R
Intermediate sec. school	34	29	.13	37	32	.22**	32	36		42	40	
Upper secondary school	23	21		21	11	.23	29	22		21	19	
No vocational degree	22	24		17	38	-.33*	15	23	-.28	14	18	
Degree below university	58	61		64	54		60	58		66	63	
University	11	11		10	4		15	10		10	11	
Income and wealth												
Monthly earnings in EUR	1815	1808	.0001**	832	721		1737	1783		912	866	
Net family income	2148	2029		2048	1970		2225	2214		2263	1999	.0001
Owner of home / flat	.34	.34		.43	.29	.16*	.42	.36		.50	.40	
Past and current employment status (in years)												
Full time work	8.4	10		5.5	6.0		7.3	8.1		5.9	5.5	
Part time work	.22	.16		1.3	1.3		.21	.17		1.4	1.3	
Unemployment	.21	.32		.24	.31		.16	.20		.21	.16	
Current employment status (in %)												
Out of labour force	1	1	-	23	34		0	1	-	21	24	
Unemployed	4	5	-.34	5	5		3	2		2	2	
Part time employed	2	1	-	21	18		3	1	-	22	19	
Full time employed	82	85		45	40		80	85		45	44	
Weekly hours	34	36	-.006	21	19		33	35		21	20	
Information on current employer (coded 0 if not employed; in %)												
Public sector	18	12	.25	18	11		25	18		19	15	
Firm size < 20	17	20		17	14		16	18		18	16	
Firm size > 2000	21	23		11	10		28	23		13	11	

Table 2.2 to be continued.

Table 2.2 continued ...

	Sports activity before											
	Less than monthly						At least monthly					
	Men			Women			Men			Women		
	Mean in subsample		Pro-bit	Mean in subsample		Pro-bit	Mean in subsample		Pro-bit	Mean in subsample		Pro-bit
Characteristics	Sport	No S.	S-NS	Sport	No S.	S-NS	Sport	No S.	S-NS	Sport	No S.	S-NS
(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)
Information on current occupation (coded 0 if not employed)												
In vocational training	.06	.05		.07	.04	-	.09	.09		.06	.05	
Self-employed	.04	.06		.03	.03	-	.03	.04		.05	.03	
Civil servant ('Beamter')	.08	.06		.02	.04	-	.15	.09		.05	.04	
Occupation: Production	.37	.42		.10	.14		.27	.38		.04	.09	
Technical	.07	.06	-	.02	.02	-	.06	.07	-.39*	.02	.02	-
Services	.30	.28		.43	.31		.38	.30		.45	.37	
Office	.19	.13		.21	.13		.27	.19		.25	.28	-.32*
Occ. with low autonomy	.17	.23		.11	.20		.09	.17		.04	.10	-.66**
… below medium auton.	.24	.26	R	.17	.14	R	.18	.25	R	.15	.18	R
… medium autonomy	.19	.17		.23	.15		.27	.18	.51**	.32	.21	.38*
… high autonomy	.17	.16		.08	.05		.20	.15	.36*	.09	.08	
… fits vocational degree	.43	.38		.34	.21		.46	.41		.37	.32	
Job prestige (Treimann, 13-78, 78: highest)	37	35		36	31		35	36	-.006*	38	37	
Health and smoking												
Satisfac. with health high	.30	.26		.23	.25	-.20*	.26	.27		.26	.25	
Satisf. w. health highest	.40	.38		.37	.34		.46	.46		.43	.39	
Visits of MD last 3 mo.	1.5	1.7	-.02	2.8	2.6		1.9	1.6		2.7	2.6	
Chronical illness	.11	.11		.17	.16		.11	.11		16	11	.28*
Days absent from work last year	4.1	4.6		3.4	3.4		4.0	4.1		2.7	2.8	
Never smoked	.43	.38		.55	.54		.49	.40	.17*	.55	.55	
General satisfaction with life (in %)												
Medium	36	41	-.27*	34	38		35	36		31	40	
High	28	28	-.24	26	26	.27	31	28	.33	33	28	
Highest	29	25		33	29	.31*	29	29		29	24	
Regional information												
Unemployment (in %)	7.7	8.1		7.9	7.8		7.8	7.2	.06**	8.3	7.8	.05*
Inhabitants per km2	16	17	-	17	17	-	17	16	-	18	16	.01**
Southern states	.39	.33	.17	.36	.37		.34	.43		.30	.37	
Central states	.16	.15		.16	.15		.17	.15		.15	.12	
Town > 500.000 inhab.	.31	.34		.30	.33		.29	.28		.34	.34	
100.000-500.000	.08	.11		.10	.11		.09	.09		.09	.10	
5.000- 20.000	.10	.09		.10	.10		.09	.11		.08	.08	
< 5.000	.07	.07		.08	.09		.06	.07		.07	.08	
City centers	.26	.29		.25	.29		.24	.24		.27	.30	-.31
# of obs; Efron's R2 in %	482	1545	9	448	1790	14	1132	339	10	653	262	15

Note: The 'no-sports sample' consists of individuals with less than monthly participation in sports activities in the year before their decision is analysed. The sports sample is made up of individuals participating in sports activities more frequently. The dependent variable in the probit is a dummy variable which is one if the individual participated at least monthly in sports activities in the relevant year when the decision is analysed. Independent variables are measured prior to the dependent variable. Coefficients are only reported when significant at the 10% level. If they are significant at the 5% (1%) level, they are marked by one (two) '*'. The probit includes a constant term and a control variable for the 'sports intensity before'. Some variables in the table are not included in the estimation. They are either marked by R (reference category), or '-' (variable deleted for other reasons like too small cell size). Some groups of explanatory variables do not add up to 100% because of variables omitted, or due to missing values.

The clearest association is that for employed women who are more likely to be observed as being active. The effect of work intensity seems to be small, if existent at all. By and large the different occupational variables confirm the general finding that individuals in 'better' jobs (having more responsibilities, requiring a higher level of training, etc.) as well as individuals with jobs in the public sector are more likely to be observed to be active in sports. The association to firm size appears to be somewhat ambiguous. Most of these differences are more pronounced for women than for men.

Health is measured by several variables. There are some 'objective' health measures, such as the number of visits of a medical doctor in the last three months, degree of disability (not presented), missing days of work due to illness in the last year, or whether the individual has any chronic diseases. Furthermore, there is a measure of self-assessed satisfaction with one's own health using an 11-point scale. Although, there is evidence that subjective health status is positively associated with sports participation, the link between previous health status and sports activities is weak. This weak links becomes even more questionable, for example, by the fact that being chronically ill is positively associated with sports participation in the female sports sample. It should however be recalled that individuals who are of particularly bad health (measured by the fact that they have been hospitalized in or before the year of the decision) were removed from the sample.

Smoking is known to be a possible important factor of participation in sports (e.g. Farrel and Shields, 2002), however, in the GSOEP it is observed only from 1998. This impedes its use as a control variable, because it might have already been influenced by previous sports participation. However, in 1999, 2001, and 2002, individuals are also asked whether they

'never smoked'. This variable is included in the probit estimation.[13] The results point in the expected direction for men, since never having smoked is positively associated with participation in sports. However, for women there appears to be no such association.

Variables measuring worries (not presented) and general life satisfaction are considered as well to capture further individual traits that may influence the decision to participate. Small differences appear in the sense that the satisfaction level of participants is higher than that of non-participants (as in Becker, Klein, and Schneider, 2006). Individual height is considered as well, but there are no apparent differences (not in table). Unfortunately, weight is measured only much later so that a pre-decision BMI could not be calculated. The same is true for alcohol and tobacco consumption.

To account for regional differences, the information on the German federal states and the types of urbanization is supplemented with regional indicators reported in the special regional files of the GSOEP allowing for an extensive socio-economic characterization of the region the individual lives in. However, it is hard to detect any systematic patterns, perhaps with the expected exception that living in a city centre seems to be (weakly) negatively associated with sports participation, if at all. The probit for the sports sample also points to a surprising positive association between sports participations and the regional unemployment rate. The explanation for this empirical finding may have to do with some other regional characteristics not captured by the depth of the regional information included. However, what this factor may be is open to speculation.

---

[13] This variable relates to the past as well as to the present and is thus less influenced by current sports participation. To avoid ignoring this important selection variable, it is included despite the endogeneity problem. However, sensitivity analysis has been performed when this variable was omitted from the specification. These results indicate that none of the conclusions depend on the inclusion of this variable.

To conclude, the results confirm most of the findings that exist in the literature so far (see Section 2.1) with the some pronounced exceptions. Furthermore, considerable heterogeneity between men and women appeared. Generally, the differences in characteristics for sport participants and non-participants are more pronounced for women than for men. Therefore, it is not surprising that the Pseudo-R^2's of the probit in the two samples of women are considerably higher than in the two samples of men.

3 The effect of sports participation on labour market outcomes: Identification and estimation

3.1 Identification

The previous section showed that participation in sports activities is not a random event. Based on this analysis, comparing earnings of sports participants and non-participants is expected to result in a positive earnings effect for the sports participants simply because better educated individuals are more likely to participate in sports. Therefore, such crude comparisons lead to biases for the 'causal effects' of sports participation that have to be corrected. Such biases can be traced back to different distributions of variables related to sports participation and outcomes (e.g. earnings 16 years later). Therefore, these variables, which may or may not be observable in a particular application, are called confounding variables or confounders in the statistical literature (e.g., Rubin, 1974). The presence of observable confounders can be corrected with various econometric methods, if these confounding variables are not affected by sports participation, i.e. if they are exogenous in this sense. Again, the previous section showed how the emphasis on particular subsamples with the same sports

status prior to the sports participation 'decision' analysed mitigates or even removes the potential endogeneity problem.[14]

The next step is to identify the variables that should be considered as confounding. The first source for such variables is the empirical literature discussed above that points to a couple of variables, which almost all are covered in our data base more detailed than in those studies. The variables in this list that are problematic in the GSOEP are life-style related variables measuring eating and drinking habits. They are measured in the GSOEP, but only in recent years. Thus, they cannot be used directly, because due to the later measurement they are very likely to be affected by previous sports participation, i.e. they are not exogenous. The literature (e.g. Farrel and Shields, 2002) suggests that drinking may in fact be related to higher sports participation and could also be negatively related to earnings. Thus, a downward bias appears to be likely. On the other hand, excess weight is related to lower sports participation and lower labour market outcomes which leads to an upward bias. There are several reasons why these biases might not be too severe: First, the missing life-style variables are correlated with other socio-economic variables that are controlled for, in particular labour market histories, earnings, type of occupation, and education, among others. Second, the biases plausibly go in different directions so some of them are likely to cancel. Third, it is reassuring that no significant effect of sports participation could be detected when treating weight, drinking and smoking formally as outcome variables in the estimation process.[15]

An alternative route to analyze the selection problem is to consider sports participation from a rational choice perspective comparing expected costs and benefits from this activity

---

[14] A remaining problem could be that people anticipate that they will start sports activities next year and change behaviour already today in anticipation of that. However, such long-term planning for a leisure activity seems to be unlikely.

[15] The exceptions to this finding are some subgroups of men for which a weight reduction can be detected.

(see for example Cawley, 2004, who used this approach to analyze eating and drinking be-haviour). The expected cost consists of direct monetary costs (e.g. buying equipment, fees for fitness studio, travel expenses to sports facilities, injuries costs), as well as foregone earnings, forgone home production, and foregone utility from other leisure activities (assuming that sports activity is a substitute for work or leisure, or both). Some types of (unpleasant) sports activities may also be associated with a direct disutility. The gains of leisure sports comes as direct utility from sports activities (fun, relaxing after an exhausting working day, etc.), as well as from the role of sports as an investment in so-called health capital. The latter can be seen as a part of an individuals' human capital as it enhances productivity and the value of leisure (see Grossmann, 1972).

What implications do these issues have for the variables that are required as controls for the empirical analysis to have a causal meaning? In fact, they are the same variables as already discussed. For example, direct costs depend on location, because sports participation is typically more expensive when living in inner cities than in suburbs or in small villages. Furthermore, opportunity costs depend on the value of the alternatives to sports, which are work, household production, and leisure. The value of these alternatives is in turn highly correlated with (and determined by) the socio-demographic variables discussed above (type of occupation, education, household composition, health, age, gender, etc.). Furthermore, their value should be related to the conditions in the local labour market. The concept of health capital appears to suggest that individuals with higher returns (or lower investment costs) should invest more in such capital. Again, it could be conjectured that the socio-demographic variables that determine the returns from work are also related to the stock of health capital. However, this remains somewhat speculative as there is not much empirical research on how to measure the returns from health capital. Furthermore, the individual discount factors should

play some role since individuals who value the future relatively more should invest more in their health capital. However, such preferences are notoriously hard to measure in survey.

The methodological approach taken to the empirical analysis in this paper can be summarized as follows: The previous section showed that some groups of individuals are more likely to participate than others. If we were able to observe all characteristics characterising these groups with different likelihoods to participate that also influence the outcomes of interest, the confounders, then we can use the fact that these variables are usually not perfect predictors for the activity levels, i.e. there are other random variations of sports participation not influencing our outcomes of interest, to compare the outcomes of members of the same group with different sports participation statuses. Obviously, for such an approach to lead to reliable results, it is crucial that all important variables jointly influencing outcomes and sports activities are observable in the data. It follows from these considerations that using the homogenous initial sample approach allows conditioning on most of the relevant exogenous variables. Thus, it will most likely remove (most of) the selection bias and does not require further restrictive statistical modelling assumptions about the relation of the outcomes, the confounders, and sports activity.

3.2   Estimation methods

As explained above, the identification and estimation problem can be tackled using an approach that exploits the panel structure of the data by performing the analysis in subsamples defined by the sports activities in the previous year and then analyzing the effects of the movements in or out of sports. In principle, once the data have been reconfigured to correspond to such a set-up, a linear or non-linear regression analysis could be used with future labour market and other outcomes as dependent variables and sports participation as well as all the other control variables as independent variables (measured in the last period when all individuals are in the same state). Such methods are well known and have been heavily used,

but they suffer from potential biases when the implied functional form assumptions are not satisfied. This is particularly worrying as these assumptions in turn imply that the effects have to be homogeneous in the population or specific subpopulation (see for example Heckman, Smith, and LaLonde, 1999). Such assumptions are clearly not attractive in this context. Recently, a flexible semiparametric method that circumvents these problems became very popular in labour economics, i.e. the method of matching (see Imbens, 2004, for a survey). It is briefly described and applied below.

Before getting into any more details, it is worth pointing out how all possible parametric, semi- and nonparametric estimators of (causal) effects that allow for heterogeneous effects are implicitly or explicitly built on the principle that for finding the effects of being in one state instead of the other (here sports activity versus no sports activity), outcomes from observations from both states with the same distribution of relevant characteristics should be compared. As discussed above, characteristics are *relevant* if they jointly influence selection and outcomes. Here, an adjusted propensity score matching estimator is used to produce such comparisons. These estimators define 'similarity' of these two groups in terms of the probability to be observed in one or the other state conditional on the confounders. This conditional probability is called the propensity score (see Rosenbaum and Rubin, 1983, for the basic ideas). A clear advantage of the class of estimators discussed in literature in this case is that they are semiparametric and allow for arbitrary individual effect heterogeneity. To obtain estimates of the conditional choice probabilities (the so-called propensity scores) used in the selection correction mechanism to form the comparison groups, the probit models presented in the previous section are applied.

The matching procedure actually used incorporates the improvements suggested by Lechner, Miquel, and Wunsch (2005). These improvements tackle two issues: (i) To allow for higher precision when many 'good' comparison observations are available, they incorporate

the idea of calliper or radius matching (e.g. Dehejia and Wahba, 2002) into the standard algorithm used for example by Gerfin and Lechner (2002). (ii) Furthermore, matching quality is increased by exploiting the fact that appropriately weighted regressions that use the sampling weights from matching have the so-called double robustness property. This property implies that the estimator remains consistent if either the matching step is based on a correctly specified selection model, or the regression model is correctly specified (e.g. Rubin, 1979; Joffe, Ten Have, Feldman, and Kimmel, 2004). Moreover, this procedure should reduce small sample as well as asymptotic bias of matching estimators (see Abadie and Imbens, 2006a) and thus increase robustness of the estimator. The matching protocol is shown in Table B.1 in Appendix B. See Lechner, Miquel, and Wunsch (2005) for more information on this estimator.

There is an issue here on how to draw inference for this rather involved estimator that is a combination of weighted radius matching and weighted regression. Although Abadie and Imbens (2006b) show that the 'standard' matching estimator is not smooth enough and, therefore, bootstrap based inference is not valid, the version of the estimator implemented here is by construction much smoother than the estimator studied by Abadie and Imbens (2006b). Therefore, it is presumed that the bootstrap is valid. The bootstrap has the further advantage in that it allows the direct incorporation of the dependency between observations generated by the specific sampling design in which some individuals may appear as several observations due to the pooling of decision windows. It is implemented following MacKinnon (2006) by bootstrapping the p-values of the t-statistic directly based on symmetric confidence intervals (rejection regions). The p-values for the non-symmetric confidence intervals are typically smaller (and some are reported in the internet appendix). Bootstrapping the p-values directly as compared to bootstrapping the distribution of the effects or the standard errors has advantages because the t-statistics on which the p-values are based are asymptotically pivotal whereas the standard errors or the coefficient estimates are not.

3.3 Alternatives for identification and estimation

One of the alternatives to the proposed approach is fixed effects panel data model. They appear to be attractive at first sight because they allow for some unobserved heterogeneity related to the selection process.[16] However, these models rely on assumptions that are unattractive in this context. First, generally, only the linear version of the fixed effects models identifies the required effects. As many of the outcome variables are binary, this is clearly unattractive. Second, the assumption of strict exogeneity of the time varying control variables used in the estimation (i.e. the assumption that the part of last years' outcome measurement not explained by the regressors does not influence next years' measurement of the regressors) is very unlikely to hold. Third, the key assumptions that the fixed effect, i.e. the part of the error that is allowed to be correlated with the regressors and captures potentially unobservable confounders, has a constant effect on the outcomes over more than 16 years would be very hard to justify in this context. A further alternative to identify the effects would be to use an instrumental variable approach (e.g. Imbens and Angrist, 1994). Such an approach requires an exogenous variable that influences the outcomes under consideration only by influencing sports participation (any direct effect is ruled out). In the present context such a variable does not appear to be available.

4 Results

4.1 Introductory remarks

Below, the effects of sports participation on various outcome measures are presented. The outcomes considered relate to success in the labour market, like earnings, wages, and

---

[16] The comparison made here is made for fixed effects models, as random effects models require strictly stronger assumptions than the methods proposed below, because random effects models do not allow for any unobservables to be correlated with the regressors (see Lechner, Lollivier, and Magnac, 2008).

employment status, as well as to various objective and subjective health measures, additional socio-demographic outcomes, and a direct measure of satisfaction with life in general. For each group of outcome variables, only a few specific variables are presented for the sake of brevity. Results for additional outcome variables are available in the internet appendix. As before, the four decision years with respect to sports participation status (1985, 1986, 1988, and 1990) are pooled to increase precision. For all outcome variables the mean effects of sport participation are estimated annually over the 16 years after the respective decision year allowing some potential dynamics to be uncovered. The exceptions are some health measures that were added to the GSOEP only recently: The effects of sports on these variables could only be estimated for one point in time. Finally, the effects presented are those for the group of individuals remaining or becoming active (so-called average treatment effects on the treated). The results for the groups becoming or remaining inactive are not presented for the sake of brevity. They are in fact very similar for women. For men, the effects are qualitatively similar as well, but in several cases about 20% to 40% smaller.

To acknowledge the considerable sex specific heterogeneity in the selection process and to uncover interesting heterogeneity, sex specific results are reported. Inference is based on symmetric bootstrapped p-values based on 499 bootstrap replications as explained in Appendix B.2.

Before discussing the effects of sports participation on various outcome measures in detail, it is useful to precisely define the 'treatment', i.e. sports participation. It is the comparison of the low activity sports states (less than monthly; denoted as 'not active' below), compared to a higher level of sports activity (at least monthly; denoted as 'active'). This contrast is conditional on the pre-decision activity state that is defined in the same way that is either measured one year (for decision years 1985 and 1986) or two years earlier (for decision years 1988 and 1990 as no sports information is available for the years 1987 and 1989). The result-

26

ing strata are called 'no sports sample', and 'sports sample', respectively. In the matching estimation, the results for the two strata are averaged to increase precision.[17]

Over the 16 years for which the effects on the outcomes are estimated, there is no guarantee that the sports statuses within the two groups remain constant.[18] Using sports participation 1 to 16 years after the decision year as outcome variables shows that the activity levels narrow as individuals switch their sport status over time. However, there is still a persistent and highly significant effect of the respective sports participation in the decision year on future sports participations, which is similar in all strata (see the internet appendix for details).

4.2   Labour market effects of sports participation

The Figure 4.1 shows the earnings and wage effects of sports participation. *Monthly earnings* are measured as gross earnings in the month before the interview. *Accumulated a*verage earnings* are the average monthly earnings until the year in question. They capture the total earnings effect over time and have the additional advantage of the averages being smoother and more precise than yearly snapshots. *Wages* are computed by dividing monthly gross earnings by weekly hours (x 4.3). These variables are coded as zero when the individual is not employed. Furthermore, they are de- or inflated to year 2000 Euros to facilitate comparisons over time and entry cohorts. The figures show the mean effects over 16 years for the men and women. A symbol on the respective line indicates that this effect is significant at the 5% level.

---

[17] This is implemented by running the estimation in the strata defined by sex. Within these two strata, the selection model is fully interacted with respect to the sports status. Results by activity level are available in the internet appendix.

[18] Keeping the sports status constant over this long period would raise the endogeneity problems discussed before because time varying covariates would have to be included to correct for dynamic selection problems. Flexible selection corrections in such a dynamic framework would require dynamic treatment models of the sort discussed by Robins (1986) or Lechner (2008). However, such models are too demanding with respect to sample size to be applicable in this context.

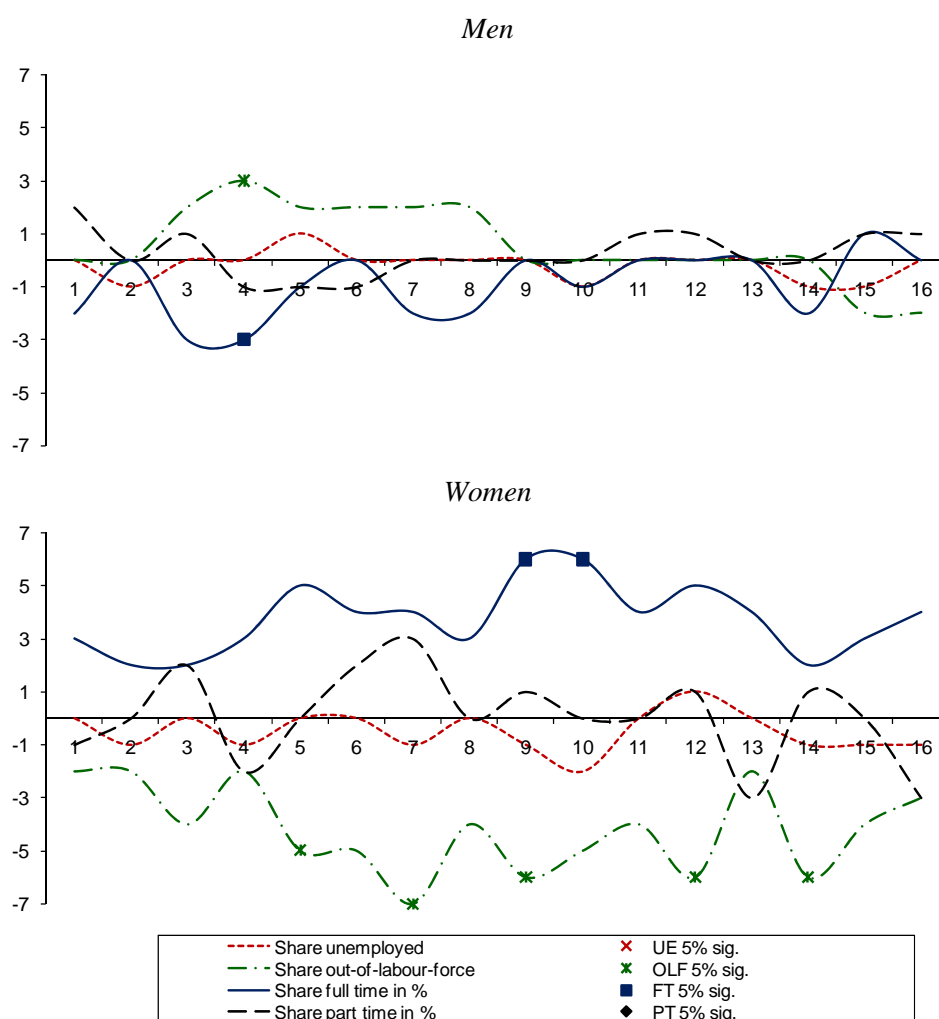*Figure 4.1: Effect of sports activity on earnings*
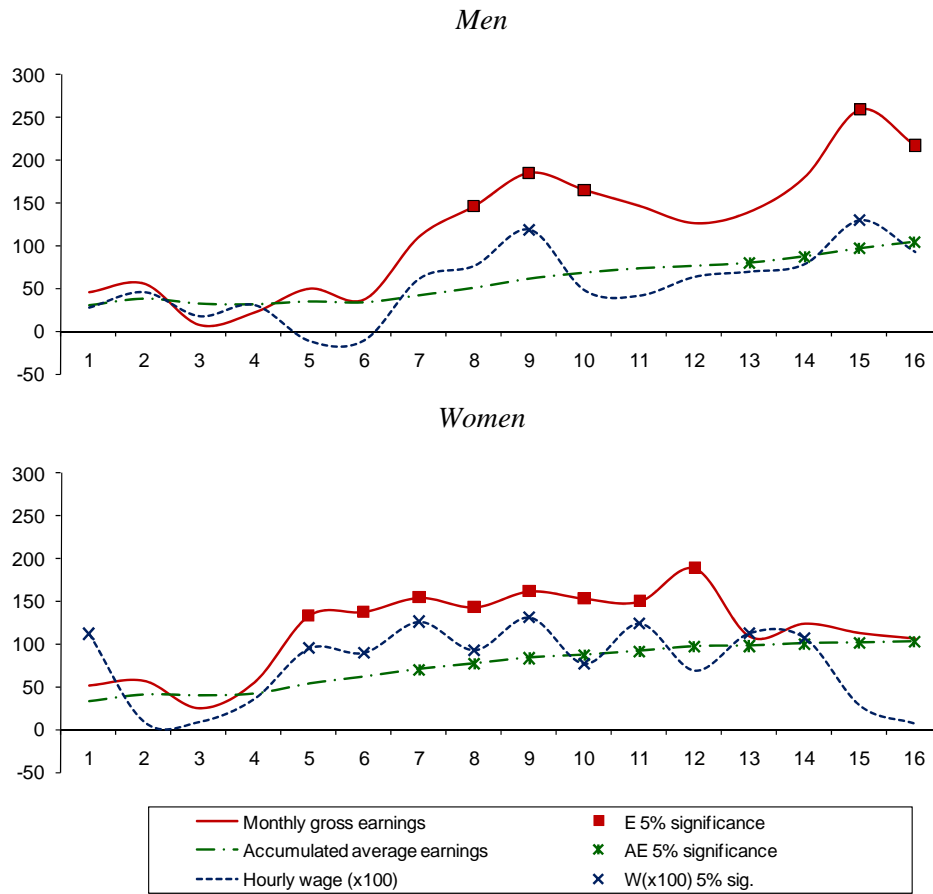
Men

Women

Note: Effects of sport participation at least monthly for the population of individuals who are active in the decision period. A symbol on the line of the mean effect indicates significance at the 5% level based on a two-sided t-test. *Monthly gross earnings* are measured as gross earnings in the month before the interview. *Accumulated average earnings* are monthly earnings summed up year by year until the year in question divided by the number the valid interviews up to the respective year. Earnings and wages are coded as zero if individuals are not employed. Wages are multiplied by 100 to be presentable on the same scale as earnings. All monetary measures are in year 2000 EUROs.

Although, estimates of the monthly earnings gains are somewhat volatile, on average after 16 years for men as well as for women there is a monthly gross earnings gain of about 100 EUR (leading to a total gain over 16 years of approximately 20.000 EUR). In most cases, these gains are at least significant at the 10% level after about 4 to 6 years (this significance level is not indicated in the figure). They appear to increase over time. Similarly, positive average wage effects of almost 1 EUR per hour are present.

Next, Figure 4.2 presents the labour supply effects of sports participation using the categories full-time work, part-time work, unemployed, and out-of-the labour force. No

significant long-run labour supply effects appear for men. However, for women there is an increase in the probability of full-time employment that goes along with a decline in the share of women considered as being out-of-the-labour force. For women, there is an increase of about 1 weekly working hours that is however rarely significant (not shown in Figure). Again, no such effect appears for men (for details see internet appendix).

Figure 4.2: Effect of sports on employment status

Note: Effects of sport participation at least monthly for the population of individuals who are active in the decision period. A symbol on the line of the mean effect indicates significance at the 5% level based on a two-sided t-test. Effects are changes in the shares of the different employment categories (in %-points).

The question arises where these positive earnings and wage effects come from, as they are not much related to differences in labour supply, at least for women. Therefore, other outcome variables are considered below that may influence productivity as well.

4.3 Other outcome measures

4.3.1 *Health effects of sports activities*

Individual health is assessed with both objective and subjective measures. Objective

measures include days spent in the hospital in the last year, the degree of disability (i.e., a

reduction in the capacity to work on a scale from 0% to 100%), the number of visits to a

medical doctor in the last three months prior to the interview, the days unable to work because

of illness in the year before the interview, as well as whether the actual case of somebody dy-

ing. These measures are supplemented by two subjective health measures: (i) individuals state

their health on a five point scale from very good to very bad (available from year 7 onwards),

and (ii) they indicate their general satisfaction with their health status on an 11-point scale.[19]

Since all health indicators show a similar pattern over time, Figure 4.3 presents only

three of them, namely the days lost at work (as a measure of direct productivity loss due to

bad health), the share of individuals reporting any disability, as well as the individually per-

ceived state of health using the five point scale (1: very good, 5: very bad). Thus, negative

values in Figures 4.3 indicate a positive health effect of sports participation. Detailed results

for the other health indicators are available in the internet appendix. The indicator of the

satisfaction with health is presented in Figure 4.4.

All in all, there are positive health effects on the subjective scale, although they are

rarely significant at the 5% level for men. Concerning satisfaction with one's own health (Fig-

ure 4.4), there is some evidence that the satisfaction increases. However, these subjective

health effects do not show up in a reduced number of lost days at work due to (temporary)

---

[19] Generally, it is considered to be no good econometric practise to use ordinal scales directly as outcome measures. However, since using (many) indicators for the specific values of the scales qualitatively leads to the same results as when using the scales directly, the effects on the ordinal scales are good summary measures in this case.

illness. However, the share of people certified as having some degree of permanently reduced work ability due to disability is decreased in the longer run. The estimate of this decrease is however volatile and only significant for women.

Figure 4.3: Effects of sports participation on health
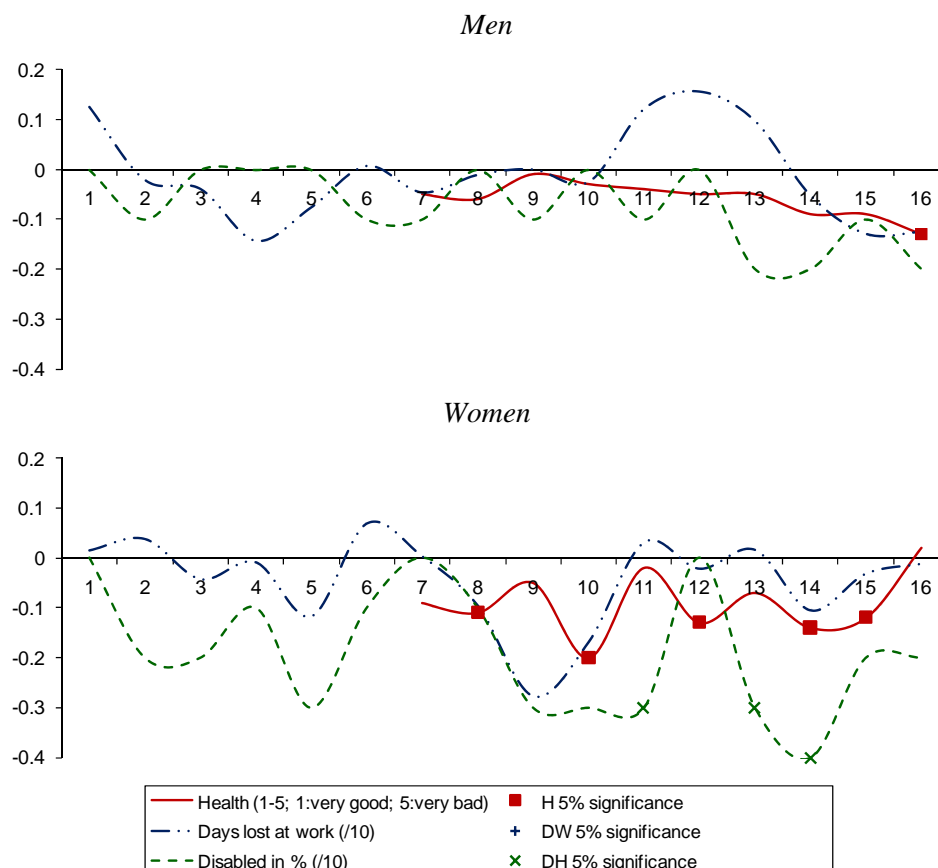
Whereas these variables are observable over a longer period, for recent years the GSOEP also contains variables describing the subjective impact of health on the tasks of daily life (see Appendix A for a detailed description) as well as alcoholic drinking behaviour and body weight. The effects on these variables, presented in Table 4.1 seem to confirm the findings for the subjective health measures. There are robust and significantly positive health effects for women and men (significance levels are indicated with '+' for 10%, '*' for 5%, and

'**' for the 1%). However, in some cases these effects are too small to be significant at conventional levels.

Table 4.1: Effects of sports participation on health (12v2) after 16 years, weight and drinking

	Men		Women	
Outcome variable	Effect	p-val. in %	Effect	p-val. in %
Mental health (summary measure)	.8	9	.9	11
Vitality	.5	42	.9	12
Social functioning	1.1*	3	.6	25
Role emotional	.6	20	.8	21
Mental health	.9+	7	1.1*	3
Physical health (summary measure)	.8+	8	.6	20
Role physical	1.1*	1	.7	21
Physical functioning	.9+	9	1.3**	0
Bodily pain	.3	56	.6	22
General health	1.4*	1	.3	61
Weight (in kg)	-1.8*	3	-.34	52
Never drinking alcohol	-.01	88	-.04	43

Note: The health measures are based on a standardized scale from 0 to 100 with standard deviation 10. 100 denotes the best and 0 the worst health status. See Appendix A.1 for details. One (two) '*' denotes significance at the 5% (1%) + denotes significance at the 10% level. Significance levels are based on a two-sided t-test. Drinking is measured on a four point scale (4: never, …, 1 regularly).

With respect to weight, there is a significant weight reduction for men of almost 2 kg, but no significant effect for women. With respect to drinking alcohol, there is no significant effect, neither for men nor for women.[20]

4.3.2   Effects of sports participation on worries, and life satisfaction, and marital status

The next step in this empirical analysis goes beyond the direct health indicators and considers general well-being measures. Three measure are presented in Figure 4.4 that should indicate different aspects of the quality of life, namely whether the individual is worried about the economic situation, his/her general satisfaction with life (ten point scale; 0: very low, 10: very high), as well as general satisfaction with health (already discussed). Additional indicators are available in the internet appendix.

---

[20] However, pre-decision weight and drinking behaviour were not available as control variables. This fact renders the results for these variables less reliable.

In both samples there is some evidence that worries about the economy in general are reduced, although estimates are volatile and significance levels vary. For men, there is also some indication that satisfaction with life in general is significantly increased in the long run, whereas for the women the effect goes in the same direction (with the exception of the last period), but appears to be too small and too noisy to become significant.

Figure 4.4: Effects of sports participation on satisfaction with life and health and worries about the economy
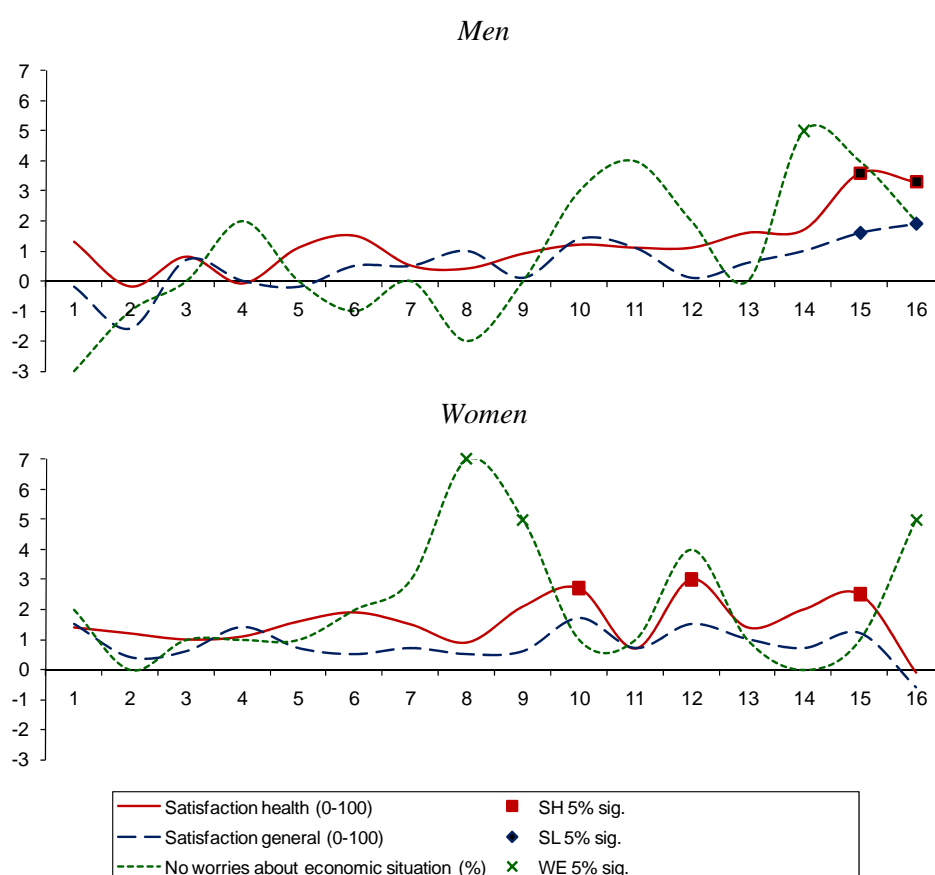
Note: Effects of sport participation at least monthly for individuals who are active in the decision period. A symbol on the line of the mean effect indicates significance at the 5% level based on a two-sided t-test.

Several variables are used to indicate marital status as well as health. Although, scattered effects show up, it is hard to detect any systematic pattern. Therefore, for the sake of brevity, these results are relegated to the internet appendix.

4.4 On the channels creating the earnings effects

One might speculate on the channel by which the gains in wage and earnings are transmitted. One channel could be health, i.e. gains in earnings just reflect the increased productivity due to better health. To check that possibility, various long-run health variables are included in the analysis as additional control variables. If the effects originate from the health effects only, then it is expected that conditional on health, the effects will disappear. Doing so reduces the long-run effects for men and women by about 15% to 20%.

When we condition in addition on general life satisfaction, worries, number of kids, and family status, then for women the earnings effects are halved. However, for men the effects are only reduced by a further 20%. These results suggest that although health and other subjective variables contribute substantially to the effects of sports activity, there remains a large unobserved and unexplained component, which is more important for men than for women. Thus, other channels, perhaps relating to social networking, are relevant as well.

4.5 Sensitivity checks

Several checks are performed to better understand the sensitivity of the results with respect to arbitrary specification and variable choices as well as to discover further important heterogeneity.

The first set of checks concerns socio-demographic variables influencing outcomes and selection that do not come as a surprise but can be planned or anticipated. Thus, the individual takes into account events that materialize in these variables one or two years ahead. If this is true, these future values should be included in the probits or sample selection rules as they indicate current or past decisions that have not yet materialized. Here, children and being married (two years ahead) are included in the probits. Furthermore, individuals with days in the hospital in the current and the following year (year 1) were removed from the sample. However, the results are robust with respect to both of these changes. In a similar attempt

several ways to specify the various health variables (different functional forms, different sets of variables) are explored, but the final results are not sensitive to different (reasonable) ways to measure health. The health variables are also used to select the sample in different ways, but again no sensitivity was detected.

The second set of checks concerns the definition of the sports participation variable. The following checks are performed: (i) Comparing the two most extreme categories (1 & 2) to the no-sports (4) category; (ii) comparing (1) to (3 & 4); (iii) comparing (2 & 3) with (4) motivated by the consideration that too much sports may be not good either and (iv) comparing (1 & 2 & 3) with (4). However, these changes did not change the results much, although it should be noted that the sharper definitions (i) to (iii) reduce the number of observations and thus leads to noisier estimates. In another check, estimation was conducted without conditioning on the sports status before (i.e. removing the interaction terms in the probit estimation). This results in more precise estimates of the effects. In particular more health variables become significant (in the expected direction). Nevertheless, this specification remains dubious because of the endogeneity problem discussed above.

To understand the robustness with respect to enforcing the balanced panel structure (required for meaningful interpretation of many of the outcome variables), the effect of sports participation on being in the balanced part of the sample has been estimated in an unbalanced panel design. It turned out that there is no such effect and thus it appears innocuous in this particular application to require a balanced panel over such a long horizon.

The age restriction may also be of concern as some fairly young individuals are included when requiring a lower age limit of 18 year, some of them may still be in the education system. Restricting the sample to individuals 24 years old and older leads to an efficiency loss due to the smaller sample, but otherwise similar results. Increasing the upper age limit to 50 instead of 44 increases precision but some of the individuals are now 65 at the end of the fol-

low-up period. Therefore, more observations withdraw from the labour market. Thus, it is much harder to detect any earnings effects.

Furthermore, the sample has been restricted to those working full-time in the relevant period to get the 'pure' earnings effects. The results point in the same direction as those for the overall sample. However, the samples are reduced considerably and the additional noise made it very hard to obtain enough precision to obtain significant estimates.

In conclusion, the results appear to be robust to reasonable deviations from the specifications underlying the conclusions drawn from Tables 2.2, 4.1, and 4.2.

5 Conclusion

This microeconometric study described the correlates of sports participation and analyzed the effects of participation in sports on long-term labour market variables, on socio-demographic variables, as well as on health and subjective well-being outcomes for West Germany using individual data from the German Socio-economic Panel study (GSOEP) 1984 to 2006. The issue that people choose their level of sports activities and, thus, participants in sports may not be comparable to individuals not active in sports, is approached by using very informative data, flexible semiparametric estimation methods, and a specific use of the panel dimension of the GSOEP.

The analysis of the selection process into leisure sports activities suggests that sports activities are higher for men than for women, and much lower for non-Germans, particularly for non-German women. Activities increase with education, earnings, and 'job quality'. Marriage, children, and older age are associated with lower sports activities.

The analysis of the effects of sports activities on outcomes revealed sizeable labour market effects. As a rough estimate, active sports increases earning by about 1.200 EUR p.a.

over a 16 year period compared to no or very low sports activities. These results translate into a rate of return on sports activities in the range from 5% to 10%, suggesting similar magnitudes than for one additional year of schooling. Increased health and improved well-being in general seem to be relevant channels to foster these gains in earnings.

Future research should focus on improving data quality in longitudinal studies to better understand how the channel from sports participation to labour market outcomes. Such improved data should include not only much more detailed health and life style data, but also more information on the intensity and type of sports activity. It would also be important to increase the sample sizes available for such studies, as the current analysis was frequently confronted with the problem that sample sizes were too small to investigate interesting heterogeneity issues. Apparently, even if such a database was initiated now, it would take a long time before it could be used for any empirical analysis. Until then, it is hoped that this paper provides valuable information about the effects of leisure sports participation on labour market and socio-demographic outcomes.

References

Abadie, A., and G. W. Imbens (2006a): "Large Sample Properties of Matching Estimators for Average Treatment Effects", *Econometrica*, 74, 235-267.

Abadie, A., and G. W. Imbens (2006b): "On the Failure of the Bootstrap for Matching Estimators", mimeo.

Aguilera, V., and M. Bernabé (2005): "The Impact of Social Capital on the Earnings of Puerto Rican Migrants," *The Sociological Quarterly*, 46, 569-592.

Andersen, H., A. Mühlbacher, M. Nübling, J. Schupp, and G. G. Wagner (2007): "Computation of Standard Values for Physical and Mental Health Scale Scores Using the SOEP Version of SF-12v2", *Schmollers Jahrbuch* 127, 171-182.

Andreyeva, T., P. Michaud, and A. van Soest (2005): "Obesity and Health in Europeans Aged 50 and above", Working Paper, Rand, 331.

Barron, J. M., B. T. Ewing, and G. R. Waddell (2000): "The Effects of High School Athletic Participation on Education and Labor Market Outcomes", *The Review of Economics and Statistics*, 82, 409-421.

Becker, S., T. Klein, and S. Schneider (2006): "Sportaktivität in Deutschland im 10-Jahres Vergleich", *Deutsche Zeitschrift für Sportmedizin*, 57, 226-232.

Bleich, S., D. Cutler, C. Murray, and A. Adams (2007): "Why Is The Developed World Obese?", NBER Working Paper 12954.

Breuer, C. (2004): "Zur Dynamik der Sportnachfrage", *Sport und Gesellschaft*, 1, 50-72.

Cawley, J. (2004): "An Economic Framework for Understanding Physical Activity and Eating Behaviors", *American Journal of Preventive Medicine*, 27 (3S), 117–125.

Cornelissen, T., and C. Pfeifer (2007): "The Impact of Participation in Sports on Educational Attainment: New Evidence from Germany," IZA DP 3160.

Crossley, Th. F., and S. Kennedy (2002): "The reliability of self-assessed health status," *Journal of Health Economics* 21 (2002) 643–658.

Dehejia, R. H., and S. Wahba (2002): "Propensity-Score-Matching Methods for Nonexperimental Causal Studies", *Review of Economics and Statistics*, 84, 151-161.

Deutscher Bundestag (2006): "11. Sportbericht der Bundesregierung," *Drucksache des Deutschen Bundstags*, 16/3750, 4.12.2006, Berlin.

Eccles, J. S., B. L. Barber, M. Stone, and J. Hunt (2003): "Extracurricular Activities and Adolescent Development", *Journal of Social Issues*, 59, 865-889.

Ewing, B. T. (1998): "Athletes and work", *Economics Letters*, 59,113–117.

Ewing, B. T. (2007): "The Labor Market Effects of High School Athletic Participation: Evidence From Wage and Fringe Benefit Differentials", *Journal of Sports Economics*, 8, 255-265.

Farrell, L., and M. A. Shields (2002): "Investigating the economic and demographic determinants of sporting participation in England", *Journal of the Royal Statistical Society A*, 165, 335-348.

Gerfin, M., and M. Lechner (2002): "A Microeconometric Evaluation of the Swiss Active Labor Market Policy," *The Economic Journal*, 112, 854-893.

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

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

Grossman, M. (1972): "On the Concept of Health Capital and the Demand for Health", *The Journal of Political Economy*, 80, 223-255.

Imbens, G. W. (2004): "Nonparametric Estimation of Average Treatment Effects Under Exogeneity: A Review", *The Review of Economics and Statistics*, 86, 4-29.

Imbens, G. W., and J. D. Angrist (1994): "Identification and Estimation of Local Average Treatment Effects," *Econometrica*, 62, 467-475.

Hollmann, W., R. Rost, H. Liesen, B. Doufaux, H. Heck, A. Mader (1981): "Assessment of different forms of physical activity with respect to preventive and rehabilitative cardiology", *International Journal of Sports Medicine*, 2, 67.

Joffe, M. M., T. R. Ten Have, H. I. Feldman, and S. Kimmel (2004): "Model Selection, Confounder Control, and Marginal Structural Models", *The American Statistician*, 58-4, 272-279.

Heckman, J. J., R. LaLonde, and J. A. Smith (1999): "The Economics and Econometrics of Active Labor Market Programs", in: O. Ashenfelter and D. Card (eds.), *Handbook of Labour Economics*, Vol. 3, 1865-2097, Amsterdam: North-Holland.

Krueger, A. B., and D. A. Schkade (2007): "The Reliability of Subjective Well-Being Measures", NBER Working Paper, 13027.

Lakdawalla, D., and T. Philipson. 2007. "Labor Supply and Weight.", *Journal of Human Resources* 42(1), 85–116.

Lechner, M. (2008): "Sequential Causal Models for the Evaluation of Labor Market Programs", forthcoming in the *Journal of Business & Economic Statistics*.

Lechner, M., R. Miquel, and C. Wunsch (2005): "Long-Run Effects of Public Sector Sponsored Training in West Germany", CEPR Discussion Paper 4851.

Lechner, M., S. Lollivier, and T. Magnac (2008): "Parametric Binary Choice models", forthcoming in P. Sevestre and L. Matyas (eds.), *The Econometrics of Panel Data*, 2nd edition, chapter 7.

Long, J. E., and S. B. Caudill (2001): "The Impact of Participation in Intercollegiate Athletics on Income and Graduation", *The Review of Economics and Statistics*, 73, 525-531.

Lüschen, G., T. Abel, W. Cockerham, and G. Kunz (1993): "Kausalbeziehungen und sozio-kulturelle Kontexte zwischen Sport und Gesundheit", *Sportwissenschaft*, 23, 175-186.

MacKinnon J. G. (2006): Bootstrap Methods in Econometrics, mimeo.

Manski, C. F., and S. R. Lerman (1977): "The Estimation of Choice Probabilities from Choice Based Samples *Econometrica*, 45, 1977-1988

Michaud, P., A. H. O. van Soest, and T. Andreyeva (2007): "Cross-Country Variation in Obesity Patterns among Older American and Europeans", *Forum for Health Economics & Policy*, 10 (2), Article 8, 1-30.

Persico, N., A. Postlewaite, and D. Silverman (2004): "The Effect of Adolescent Experience on Labor Market Outcomes: The Case of Height", *Journal of Political Economy*, 112, 1019-1053.

Prentice, A. M., and S. A. Jebb (1995): "Obesity in Britain: gluttony or sloth", *British Medical Journal*, 311, 437-439.

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

Robins, J. M. (1986): "A New Approach to Causal Inference in Mortality Studies with Sustained Exposure Periods - Application to Control of the Healthy Worker Survivor Effect", *Mathematical Modelling*, 7, 1393-1512.

Rosenbaum, P., and D. Rubin (1983): "The Central Role of the Propensity Score in Observational Studies for Causal Effects", *Biometrica*, 70, 41-55.

Rubin, D. B. (1974): "Estimating Causal Effects of Treatments in Randomized and Nonrandomized Studies", *Journal of Educational Psychology*, 66, 688-701.

Rubin, D. B. (1979): "Using Multivariate Matched Sampling and Regression Adjustment to Control Bias in Observational Studies", *Journal of the American Statistical Association*, 74, 318-328.

Ruhm, C. J. (2000): "Are Recessions Good For Your Health?", *The Quarterly Journal of Economics*, 617-650.

Ruhm, C. J. (2007): "Current and Future Prelevence of Obesity and Severe Obesity in the United States", *Forum for Health Economics & Policy*, 10 (2), Article 6, 1-26.

Sabo, D., K. E. Miller, M. J. Melnick, M. P. Farrell, and G. M. Barnes (2005): "High School Athletic Participation And Adolescent Suicide: A Nationwide US Study", *International Review For The Sociology of Sport*, 40/1, 5–23.

Schneider, S., and S. Becker (2005)**:** "Prevalence of physical activity among the working population and correlation with work-related factors. Results from the First German National Health Survey", *Journal of Occupational Health*, 47, 414-423**.**

Statistisches Bundesamt (2005), "Körperliche Aktivität", *Robert-Koch-Institut, Gesundheitsberichterstattung des Bundes*, Heft 26.

Stevenson, B. A. (2006): "Beyond the Classroom: Using Title IX to Measure the Return to High School Sports", *American Law & Economics Association Annual Meetings*, Year 2006, Paper 34.

US Department of Health and Human Services, Centers for Disease Control and Prevention and National Center for Chronic Disease Prevention and Health Promotion (1996): "Physical Activity and Health: A Report of the Surgeon General", *International Medical Publishing*, Atlanta, 87-144.

Wagner, G. G., J. R. Frick, and J. Schupp (2007), "The German Socio-Economic Panel Study (SOEP) –Scope, Evolution and Enhancements", *Schmollers Jahrbuch*, 127, 139-169.

Wellman, N. S., and B. Friedberg (2002): "Causes and consequences of adult obesity: health, social and economic impacts in the United States", *Asia Pacific Journal of Clinical Nutrition*, 11 (Suppl): S705–S709.

Weiss, O. and P. Hilscher (2003): "Wirtschaftliche Aspekte von Gesundheitssport.", *Forum Public Health*, Heft 2003/41, 29 - 31.

Appendix A: Data

A.1 Definition of some important variables

This section provides some additional information on key variables, such as the variables defining sports participation, outcomes, and factors to control for. Discussing all of the latter variables would go beyond the space constraints of this paper, so the discussion is restricted to some variables that are important as well as non-standard, such as the health information as well as further subjective indicators of the quality of life.

A.1.1 Sports participation in the GSOEP

The information on leisure sports activity differs over the years. For example, in the initial survey of 1984, the relevant question asked in three categories whether people do sports in their free time ("How often do you engage in the following activities in your free time? Active sports: never / rarely; occasionally; often / regularly"). Individuals answering 'never / rarely' and 'occasionally' constitute the *no-sports sample* with respect to the sports decision in 1985, whereas the remaining group constitutes the *sports sample*.

In 1985 and thereafter there were two types of questions. Both are more precise than the 1984 version: The first type says "Which of the following activities do you do in your free time? Please enter how often you practice each activity. … Active sports participation: each week; each month; less often; never". This question was posed in 1985, 1986, 1988, 1992, 1994, 1996, 1997, 1999, 2001, and 2005. The alternative formulation used in 1990, 1995, 1998, and 2003, was "How frequently do you do the following activities? … do sports: daily; once per week; once per months; less than once a month; never". Although, the wording is not exactly the same, once the extreme categories (daily, once a week as well as never, less than monthly) of the second type of the questions are aggregated, both types of questions appear to be sufficiently similar to be used in combination. This is also corroborated by a comparison of

the respective descriptive statistics over time (see Table 2.1. and the discussion in Section 2.3.1). A more serious problem is that for the years 1987, 1989, 1991, 1993, 2000, 2002, and 2004 no such information is available. When required for the definition of the pre-participation status and the outcomes, the missing information is taken from the previous year.

A.1.2 Health information

Health is measured by several variables. One of the health questions uses a 5-point scale and the following wording: "How would you describe your health at present? Very good; good; satisfactory; poor; very poor." Further variables for satisfaction with health are based on the following wording "How satisfied are you today with the following areas of your life? Please answer by using the following scale, in which 0 means totally unhappy and 10 means totally happy. If you are partly happy and partly not, select a number in between. How satisfied are you ... with your health?".[21]

There may be an issue with the quality of the content of the subjective health information. Although recent work suggests that the quality of self-assessed health data may have some random component that may be related to other socio-economic variables (i.e., Crossley and Kennedy, 2002), the fact that a panel data set is used that keeps these factors constant over time and that many socio-economic characteristics are conditioned on in the empirical analysis suggests that these issues are not particularly relevant for this analysis.

Nevertheless, these subjective, qualitative measure are supplemented by more objective health measure as the number of doctor visits in the last three months, days hospitalized, degree of disability (0 to 100%), whether the individual experiences any chronicle diseases, as well as the number of days unable to work in the last year. All of these variables are available

---

[21] All translations of the questions from the (German) questionnaires are taken from the official website of the GSOEP (http://panel.gsoep.de/soepinfo2006).

since the beginning of the survey. Therefore, they can be used to control for 'pre-sports-decision' health conditions and used as outcome variables.

In 2002, the GSOEP biannually added information based on how health status is impairing daily life. Since the measurements relate to 2002 and later, these variables do not play any role as control variables, but are used as outcome variables only. The respective questions are shown in Figure A.1.

Figure A.1: Health measured as impact on daily life (SF-12x2)

84. When you ascend stairs, i.e. go up several floors on foot:
Does your state of health affect you greatly, slightly or not at all?

Greatly ....☐
Slightly ....☐
Not at all ....☐

85. And what about having to cope with other tiring everyday tasks,
i.e. where one has to lift something heavy or where one requires agility:
Does your state of health affect you greatly, slightly or not at all?

Greatly ....☐
Slightly ....☐
Not at all ....☐

86. Please think about the last four weeks.
How often did it occur within this period of time, ...

	Always	Often	Some-times	Almost	Never
• that you felt rushed or pressed for time?	☐	☐	☐	☐	☐
• that you felt run-down and melancholy?	☐	☐	☐	☐	☐
• that you felt relaxed and well-balanced?	☐	☐	☐	☐	☐
• that you used up a lot of energy?	☐	☐	☐	☐	☐
• that you had strong physical pains?	☐	☐	☐	☐	☐

• that due to physical health problems
  – you achieved less than you wanted to at work or in everyday tasks? ☐ ☐ ☐ ☐ ☐
  – you were limited in some form at work or in everyday tasks? ☐ ☐ ☐ ☐ ☐

• that due to mental health or emotional problems
  – you achieved less than you wanted to at work or in everyday tasks? ☐ ☐ ☐ ☐ ☐
  – you carried out your work or everyday tasks less theroughly than usual? ☐ ☐ ☐ ☐ ☐

• that due to physical or mental health problems you were limited socially, i.e. in contact with friends, acquantances or relatives? ☐ ☐ ☐ ☐ ☐

Note: English translation of the 2004 GSOEP questionnaire.

The empirical analysis uses these variables, the subscales that relate to different types as well as the overall state of mental and physical health. All computed scales are normalised

to lie between 0 and 100. They are normalized for the year 2004 to have a mean of 50, and a standard deviation of 10. The technical details on how the scales are computed are described in Andersen, Mühlbacher, Nübling, Schupp, and Wagner (2007).

In addition to these variables, there is also information on body weight and height (and thus BMI) which are used as outcome variables. Furthermore, since height is (almost) time constant, it is used as control variable as well.

A.1.3 Further subjective variables

The questions about worries are phrased in the following way: "How about the following areas? Do they worry you? … general economic development: ... Very worried, slightly worried, not worried". The variable used in the empirical analysis is an indicator for 'very worried'.

Finally, the question about satisfaction with life in general is worded in the following way: "At the end we would like to ask you for your satisfaction with your entire life. Please answer by using the following scale, in which 0 means totally unhappy and 10 means totally happy. How happy are you at present with your life as a whole? …".

Of course, similar concerns as those related to the subjective health measured may be raised with regard to subjective well-being measures.[22] Again, note that this issue would only be relevant, if there was a systematic difference in the reliability between participants and nonparticipants in sports activities. It is very hard to see why this should be the case.

---

[22] However, Krueger, and Schkade (2007) study the reliability of such measures and conclude optimistically that "While reliability figures for subjective well-being measures are lower than those typically found for education, income and many other microeconomic variables, they are probably sufficiently high to support much of the research that is currently being undertaken on subjective well-being, particularly in studies where group means are compared (e.g., across activities or demographic groups)." (last sentence of their abstract).

A.2 Sample selection rules

The motivation and construction of the sports and no-sports sample, as well as the pooling of the different sport-participation decisions are already discussed in the main part of the text. The following additional sample selection rules are applied: (i) individuals without valid sports information in the relevant years of and before the participation decision are not taken into consideration. (ii) The analysis is based on a balanced panel over up to 19 years so that the long-term outcome variables as well as the covariates have meaningful measurements. Using an unbalanced panel for the 16 years in which the outcomes are measured, sports participation has no effect on the probability of being observed in the balanced part of the sample. Thus, there is no need to worry that requiring balancing does induce any substantial bias in the results presented. (iii) Individuals are restricted to be aged between 18 and 44. The lower age limit is to avoid analyzing individuals still in school, whereas the upper limit is imposed to avoid that retirement issues become too important, as individuals will not be older than 60 when their long-term outcomes are measured. Fourth, only individuals not disabled in the years of and before the participation decision are considered. Furthermore it is required that during the year of the decision as well as the year after the decision the individual must not have stayed in a hospital. Both restrictions are imposed to be able to concentrate on the healthy part of the population. (iv) due to very small cell sizes, individuals in agriculture and mining, etc., both physically demanding occupations, are removed.

Appendix B: Further information on the econometric methods used

B.1 Details of the matching estimator

For the sake of completeness, the matching protocol for the estimator used here is reproduced below. For further details the reader is referred to Lechner, Miquel, and Wunsch (2005). The role of the matching step in the estimation procedure is explained in Section 3.2.

Table B.1: Matching protocol for the estimation of the average effect for sports participants

Step 1	Estimate a probit model to obtain the choice probabilities conditional on covariates for all observations: $\hat{P}(X_i)$
Step 2	Restrict sample to common support: Delete all observations with probabilities larger than the smallest maximum and smaller than the largest minimum of all subsamples defined by S. In each of the 4 samples no more than 20 observations had to be removed.
Step 3	*Estimate the respective (counterfactual) expectations of the outcome variables.* The following steps are performed: **Standard propensity score matching step (binary treatments)** a-1) Choose one observation in the subsample defined by participation in sports and delete it from that pool. b-1) Find an observation in the subsample of non-participants that is as close as possible to the one chosen in step a-1) in terms of $\left[\hat{P}(x), \tilde{x}\right]$. 'Closeness' is based on the Mahalanobis distance. Do not remove that observation, so that it can be used again. c-1) Repeat a-1) and b-1) until no participant in sports is left. **Exploit thick support of X to increase efficiency (radius matching step)** d-1) Compute the maximum distance (d) obtained for any comparison between treated and matched comparison observations. a-2) Repeat a-1). b-2) Repeat b-1). If possible, find other observations in the subsample of non-participants in sports that are at least as close as $R \cdot d$ to the one chosen in step a-2) (to gain efficiency); we choose R to be 90%. Do not remove these observations, so that they can be used again. Compute weights for all chosen comparisons observations that are proportional to their distance (calculated in b-1). Normalise the weights such that they add to one. c-2) Repeat a-2) and b-2) until no participant in sports is left. d-2) For any potential comparison observation, add the weights obtained in a-2) and b-2). **Exploit double robustness properties to adjust small mismatches by regression** e) Using the weights $w(x_i)$ obtained in d-2), run a weighted linear regression of the outcome variable on the variables used to define the distance (and an intercept). f-1) Predict the potential outcome $y^l(x_i)$ of every observation in l (no sports) and m (*sports)* using the coefficients of this regression: $\hat{y}^l(x_i)$ f-2) Estimate the bias of the matching estimator for $E(Y^l \mid S = m)$ as: $\sum_{i=1}^{N} \dfrac{1(S=m)\hat{y}^l(x_i)}{N^m} - \dfrac{1(S=l)w_i\hat{y}^l(x_i)}{N^m}$. g) Using the weights obtained by weighted matching in d-2), compute a weighted mean of the outcome variables in *the non-active*. Subtract the bias from this estimate. **Final estimate** h) Compute the treatment effect by subtracting the weighted mean of the outcomes in the comparison group of non-active from the weighted mean in the group of sports participants.

Note: When a particular outcome variable Y is binary, binary logits estimated by weighted maximum likelihood (see Manski and Lerman, 1977) are used instead of weighted linear regressions. However, since all these regression type adjustments are post-matching and thus strictly local, using regressions or logits does not change the results in any significant way (for the binary variables).

B.2   Details of the implemented bootstrap procedure

After having obtained a 'normal' t-statistic $(\hat{t})$ for the test that the effect is zero, the bootstrap is implemented using the following steps.

1)   Draw a random (bootstrap) sample from the initial population in the GSOEP.

2) Impose all sample selection rules and pool data over the four starting periods.

3) Estimate the t-statistic for each bootstrap replication ($\{\hat{t}_r\}$)

4) Repeat 1) to 3) R times (R=499) and obtain $\{\hat{t}_1,...,\hat{t}_R\}$. As we are interested in the 5%-level of significance ($\alpha = 0.05$), 499 fulfills the criterion given by MacKinnon (2006), namely that $\alpha(R+1)$ should be equal to an integer (*100* in our case).

5) Compute the symmetric p-value as: $\hat{p}^* = \dfrac{1}{R}\sum_{r=1}^{R} \underline{I}(|\hat{t}_r| > |\hat{t}|)$. $\underline{I}(\cdot)$ denotes the indicator function which is one if its argument is true.