Childhood Sporting Activities and Adult Labour-Market Outcomes

Charlotte Cabane, Andrew E. Clark

PARIS SCHOOL OF ECONOMICS
ÉCOLE D'ÉCONOMIE DE PARIS

WORKING PAPER N° 2013 – 34

Childhood Sporting Activities and Adult Labour-Market Outcomes

Charlotte Cabane
Andrew E. Clark

JEL Codes: J24, J13, L83, I2

Keywords: Job characteristics, Education, Sport, School

PARIS-JOURDAN SCIENCES ECONOMIQUES
48, BD JOURDAN – E.N.S. – 75014 PARIS
TÉL. : 33(0) 1 43 13 63 00 – FAX : 33 (0) 1 43 13 63 10
www.pse.ens.fr

CENTRE NATIONAL DE LA RECHERCHE SCIENTIFIQUE – ECOLE DES HAUTES ETUDES EN SCIENCES SOCIALES
ÉCOLE DES PONTS PARISTECH – ECOLE NORMALE SUPÉRIEURE – INSTITUT NATIONAL DE LA RECHERCHE AGRONOMIQU

Childhood Sporting Activities and Adult Labour-Market Outcomes*

Charlotte Cabane† and Andrew E. Clark‡

October 2013

Abstract

We here ask whether sports participation at school is positively correlated with adult labour-market outcomes. There are many potential channels for this effect, although, as usual, identifying a causal relationship is difficult. We appeal to two widely-separated waves of Add Health data to map out the correlation between school sports and adult labour-market outcomes. We show that different types of school sports are associated with different types of jobs and labour-market insertion when adult. We take the issue of the endogeneity of sport seriously and use data on siblings in order to obtain estimates that are as close to unbiased as possible. Last, we compare the effect of sporting activities to that of other leisure activities.

JEL Codes: J24, J13, L83, I2.
Keywords: Job characteristics, Education, Sport, School.

*We would like to thank Wladimir Andreff, Nicolas Jacquemet, Yannick L'Horty, Rob Simmons and participants at the AFSE conference (Paris), the EALE conference (Paphos), the 2nd European Conference in Sports Economics (Paris), the Journées de Microéconomie Appliquée (Sousse), RES 2013, the TEPP Research on Health and Labour Conference (Le Mans) and the Lech Am Arlberg Labor Economics, PSE-WIP and SIMA seminars for helpful comments. This research uses data from Add Health, a program project designed by J. Richard Udry, Peter S. Bearman, and Kathleen Mullan Harris, and funded by a grant P01-HD31921 from the National Institute of Child Health and Human Development, with cooperative funding from 17 other agencies. Special acknowledgment is due Ronald R. Rindfuss and Barbara Entwisle for assistance in the original design. Persons interested in obtaining data files from Add Health should contact Add Health, Carolina Population Center, 123 W. Franklin Street, Chapel Hill, NC 27516-2524 (www.cpc.unc.edu/addhealth/contract.html). This research was supported by the UK Department for Work and Pensions, the U.S. National Institute of Aging (Grant No R01AG040640) and private donations.

†SEW, University of St. Gallen, Varnbüelstrasse 14, CH-9000 St. Gallen, Switzerland. E-mail: charlotte.cabane@unisg.ch

‡Paris School of Economics - CNRS (UMR 8545), 48 Boulevard Jourdan, 75014 Paris, France. E-mail: Andrew.Clark@ens.fr.

1 Introduction

There is by now a fairly-wide range of research highlighting that those who practise, or have practised, sport do better on many levels in their lives: the results here refer to educational outcomes (Pfeifer & Cornelißen, 2010, Felfe *et al.* , 2011, Rees & Sabia, 2010, Long & Caudill, 1991), the labour market (Long & Caudill, 1991, Rooth, 2011, Lechner, 2009, Kosteas, 2012, 2011, Barron *et al.* , 2000, Ewing, 1998) and health. There are a number of potential causal readings of this relationship. An obvious one is that playing sports helps children to develop both cognitive and non-cognitive skills. However, it is also entirely possible that the correlations found in the literature reflect a hidden common factor (the type of school perhaps, or the way in which the child has been raised by her parents) affecting both childhood sport and the outcome in question. In this case, children with greater cognitive and non-cognitive skills will indeed be more involved in sports when young, but the relationship will not be causal (in the sense that if we exogenously made children play more sport, their cognitive and non-cognitive skills would remain unchanged).

A second channel works via the effect of sport on health and fitness. Health is a significant component of human capital, and thus directly has an impact on adult productivity. The related empirical work has underlined the existence of a beauty premium (of which physical fitness might be thought to be one aspect) on the labour market (Mobius & Rosenblat, 2006, Hamermesh & Biddle, 1994).

Third, sport provides opportunities for socializing, such that those who play have larger and more diversified social networks. Playing team sports or practising sport in a club allows individuals become acquainted with others and enlarge their social circle. It has been widely-demonstrated that networks play an important role in the labour market.

Last, those who play sport arguably convey a positive image of themselves, in that sport is often considered to be associated with desirable qualities: self-confidence, persistence, fair play, team spirit, and motivation. Sporting activity is often mentioned on CVs, and is an easy subject to broach in professional contexts. It can therefore be argued to constitute a positive signal on the labour market (Lechner, 2009, Rooth, 2011).

We here consider the link between childhood sport participation and adult labour-market outcomes. Our analysis differs from that in the existing literature as we are able to appeal to rich data allowing us to link childhood sports to a number of aspects of the job the individual holds thirteen years later: we shall here concentrate on having managerial responsibilities and the freedom to make important decisions in one's job (autonomy). We are also able to distinguish between different types of sport at school

(individual or team), and compare the effect of sport to that of other social activities undertaken when at school.

We also explicitly address the issue of the endogeneity of sporting activities (whereby, for example, richer students, those at certain types of schools or those in better health may have better access to sporting facilities, and better opportunities on the labour market) in order to try to tease out a causal relationship between childhood sporting activity and adult labour-market success. One obvious approach here is to include in the regression analysis the variables which help determine whether individuals practise sport or not. We also introduce information on the individual's behaviour and ability when young in order to capture as much of the heterogeneity as possible. In addition, the structure of the Add Health data used here allows us to introduce school fixed effects into all of our regressions. Finally, we are able to use data on siblings (and twins) in order to capture any family fixed effects.

While there has been a fair amount of work on the relationship between sport and labour market outcomes, this literature has often appealed to data on college sport participation, whereas we here use sport information that comes from much earlier in life (during middle and high school).[1] Existing work also mostly does not distinguish between types of sport, and only considers one labour-market outcome (wages), whereas our focus here is rather on the type of job that the individual holds.

The remainder of the article is organized as follows. The following section reviews some of the existing literature, and then Section 3 describes the data that we use. Section 4 presents the empirical framework and the identification strategy, and section 5 discusses the results. Finally, Section 6 concludes.

2 Some Existing Literature

Long & Caudill (1991) use data from the National Collegiate Athletic Association (NCAA) on college students who are top athletes. They find a positive relationship between sports practice and educational success, with a 4% higher graduation rate for men and a 9% higher rate for women. Male college athletes are also subsequently found to earn higher wages, ten years later. They suggest three explanations for these sport premia. First, being a former college athlete sends out a positive signal on the labour market regarding ability, as in the signalling model of Spence (1973). Those who choose to play sport also have unobservable characteristics (concentration, stamina, teamwork, or something else) which are valued on the labour market. A second reading of this

[1] Which is arguably a time of life when individuals are more open to learning than when they are at college (Heckman & Kautz, 2012, Cobb-Clark & Schurer, 2012).

correlation is more causal: practising sport increases the individuals' soft skills. Here it is sport which teaches individuals to work in a team, be competitive, have self-discipline, and so on. Again, these productive soft skills are valued by firms. Last, as Long & Caudill (1991) focus on relatively well-known athletes, there may be something of a reputation effect. Firms may then hire former athletes as they provide the company with a good image (firm sponsorship of sport personalities, for example Zinedine Zidane by Danone, obviously appeals to this reputation effect on a grander scale).

Pfeifer & Cornelißen (2010) use data from the German Socio-Economic Panel (SOEP) to argue that the sport-labour market relationship is in fact causal, with high-school sports increasing productivity. Sport is argued to help children to develop self-esteem, a competitive spirit, tenacity, motivation, discipline and responsibility; these non-cognitive skills are all rewarded at school and are useful for the learning process. Sport also leads to better student health, which directly increases productivity. The results in Pfeifer & Cornelißen (2010) differ by gender, with the sport effect being larger for girls. The proposed interpretation is that boys and girls do not start out with the same non-cognitive skill endowment: girls are thought to be less competitive and to have lower self-esteem. They therefore have a relatively greater amount that can be gained from sports practice: the marginal productivity of sports, as it were, is greater for them. This gender difference is widespread in the literature.

Felfe *et al.* (2011) uncover a positive correlation between childhood sports participation and cognitive and non-cognitive skills in a cross-section of German children aged from 3 to 10 (KiGGS). This correlation is larger than that with other, more passive, leisure activities (watching TV) which are less productive in terms of child development.

Rees & Sabia (2010) analyse the relationship between academic performance and sports participation in data from the National Longitudinal Study of Adolescent Health (Add Health), which is the same data that we use here. They remain somewhat cautious about the existence of a causal relationship between the two, but suggest a positive effect of sport participation on aspirations to attend college.

Stevenson (2010) appeals to Title IX legislation (which concerned women's access to sports in the US) to measure the impact of sport participation on female college attendance and labour-force participation. A 10 percentage-point rise in State-level sport participation is associated with 0.04 more years of schooling and a 1.5 percentage point rise in the employment rate.

The research in Kosteas (2012) considers the link between frequent exercise and wages using NLSY79 data. The correlation is positive, significant and large, with wages

found to be 6 to 10% higher. Previous work in Kosteas (2011) using the same data reveals that those involved in more than one club during high school have a significantly higher probability of being a supervisor on the labour-market, although there is no evidence that sports clubs in particular are significantly linked to supervisory jobs.

Lechner (2009) analyses sport practice in SOEP data, which is found to be equivalent to one additional year of schooling in terms of labour-market outcomes. Three channels are identified. Sport participation improves both mental and physical health, which feeds through to higher productivity. However, as noted in Long & Caudill (1991) above, the relationship may also reflect the correlation between sports and unobservable individual characteristics which are valued on the labour market.

Rooth (2011) is able to clearly demonstrate the importance of sport participation as a labour-market signal, using an experimental approach (a correspondence study) in Sweden which allows the evaluation of particular individual characteristics on hiring. The inclusion of sport as a leisure activity in CVs increases the likelihood of interview, with a large effect size: sport participation has the same interview effect as an additional 1.5 years of work experience. Rooth distinguishes this hiring effect by sport type (and gender), which is relatively unusual in the literature.

While the above literature has considered the correlation between sport participation and labour-market outcomes, a related literature has attempted to set out the determinants of sporting activity itself. Children's participation to a considerable extent reflects parental preferences and support. Raudsepp (2006) suggests that fathers might act as a role model for sons, but not necessarily for daughters. Gustafson & Rhodes (2006) review the literature on the link between child and parental sports participation (covering 30 papers), and conclude that parental support plays a role in child sporting activity. The correlation between parental support in terms of transportation, payment of fees and encouragement, on the one hand, and child sport participation is strongly positive; that between child and parental sport participation less so. Farrell & Shields (2002) also underline that household sport preferences are homogeneous, and infer that sport preferences are either contagious or that people who like sports tend to match together.

Related work has then looked at parents' sports participation, both theoretically (Downward, 2007) and empirically (Farrell & Shields, 2002, Downward, 2007, Hovemann & Wicker, 2009, Breuer, 2006, Breuer & Wicker, 2008, Humphreys & Ruseski, 2011). Both types of analyses come to similar conclusions: those who participate in sport have very similar characteristics to those who are successful on the labour market. Sport depends positively on education and household income, while ethnic minorities, older

individuals and those in worse health practise sports less.[2] The determinants of sports and income which differ in sign are those related to the household and marital status.[3]

We here contribute to this literature using panel data over a relatively long period, covering schoolchildren through to their mid-late twenties. We also look at aspects of the job type which have been relatively ignored in the literature: managerial responsibilities and autonomy at work.

3 Data

We use data from the National Longitudinal Study of Adolescent Health (Add Health), which is a *"longitudinal study of a nationally representative sample of adolescents in grades 7-12 in the United States during the 1994-95 school year"*.[4] There are currently three follow-up waves available after the first in 1994-95. In the most recent 2008 wave, the individual respondents are 24 to 32 years old and provide information on their labour-force status; our dependent variables come from this last wave. The Add Health data is particularly apt for the question at hand in a number of ways. First, it provides information on a number of different types of sport participation in Wave I, when the individual was aged between 12 and 18, which we can link to their labour-market status 13 years later. In particular we are able to compare team and individual sports, as well as the effect of sport compared to other leisure activities. The data also include a wide array of objective individual information at Wave IV, including health, lifestyle and education, and a number of subjective variables. The first wave of Add Health data, from which our sport information is obtained, results from interviews with a number of different schoolchildren within the same school class (there are a total of 142 different schools in Wave I). As such, we are able to control for school fixed effects in all of our regressions. Last, the data includes a certain (smallish) number of pairs of siblings, which will be useful later when we consider endogeneity.

We focus on Add Health Wave IV respondents who have completed their education and are employed at the time of the interview. The information on labour-force status, completed education, health, and marital and labour-force status comes from Wave IV; other individual characteristics such as gender, ethnicity and so on are taken from wave I. There are over 10,000 workers in Wave IV, split equally by sex, although missing values on some variables produces somewhat smaller numbers of observations in the

[2]In addition, Anderson (2001) suggests that when ethnic minorities actually do practise sport they are more likely to over-invest in it and under-invest in education, and may thus end up worse off.

[3]The married participate less in sport, while the correlation with children depends on the individual's gender and the age of the children: young children lower adult sport participation, while the presence of children is positively correlated with male sport participation.

[4]See the Add Health webpage: http://www.cpc.unc.edu/projects/addhealth.

regressions.

Our sport information covers the frequency with which individuals practised sports, and the type of sport in which they engaged, at Wave I (when respondents were at school in grades 7 to 12). Sport is divided in up into *active transportation*, *exercise* and *active sports*. The frequency of practice of each sport per week is reported on an ordered scale: none, once or twice, between three and four times, and five or more times. we recode these values to reflect weekly frequencies of respectively 0, 1.5, 3.5 and 5.5 in order to be able to treat sport frequency as a cardinal variable.

Active transportation refers to cycling, roller-skating etc., which sports do not require any particular facilities or sporting structures (such as clubs). *Exercise* can be considered to refer to individual types of sports (the examples given in the questionnaire are jogging, dancing, gymnastics and so on): individuals most often will need some kind of sporting facilities in order to practise these sports, but do nonetheless not necessarily belong to a club. Finally, *active sports* refers directly to team sports (the examples given being basketball, baseball, soccer etc.). Participation in this last type of sport involves both being part of a club or a team and some particular sport facilities (i.e. a ground or court).

These types of sport are associated with different individual characteristics or skills. The practice of individual sports may reflect self-discipline and motivation, while individuals will need team spirit for successful participation in team sports. The Wave I respondents declare sports practice of on average over six times per week: one *active transportation*, around three individual sport activities and just over two of team sport. Female students report less sporting activity than do their male counterparts, but practise relatively more individual sports than team sports (which latter are preferred by male students). For both genders *active transportation* is the least frequent sport: we will not consider it in the remainder of the paper.[5]

As might be suspected, individual and team sport practice are correlated to some extent: over 40% of those who play no individual sport play no team sport either, for example. However, our cardinal versions of these two sport variables are only correlated at under 0.3, which suggests that it is potentially useful to examine their impact separately.

While we know how much sport our school respondents declare, we do not know when or where they do it. This could be during recess at school, or outside of school in a club. We do know which year the children are in, and can split them up into middle school or

[5]It is of course easy enough to do so. There turn out to be no significant correlations between *active transportation* and any of the adult labour-market outcomes we analyse here.

high school. Our underlying idea here is that those in middle school may be more likely to play sports during recess, whereas those in high school may undertake other activities at this time. The data do actually show lower sport participation amongst high-school children: in fact, less time is spent on all types of leisure activities in high school.

We find, as in the SOEP data in Pfeifer & Cornelißen (2010), that school performance is related to sport participation. In particular, good grades in Maths and Sciences are correlated with sport frequency: most of the children who earn A-grades in Maths and Sciences practise sport at least five times per week.[6] On the contrary, English and History grades are higher for those who play sports the least.[7]

We originally retained five different indices of adult labour-market success: having a paid job of at least 10 hours per week (which individuals are referred to as *workers*), job satisfaction, managerial responsibilities, the freedom to make important decisions in one's job, and log of annual earnings.[8] Some individuals who work under 10 hours also answer the labour-market questions: our results below refer systematically to workers. Almost two-thirds of Wave IV respondents (65.1%) are workers. Table 1 shows descriptive statistics by working status. Those who do not work are more likely to be ethnic minorities, women and have more children; they are also slightly less educated and have less-educated parents. Last, they participated less in team sports when at school in Wave I.

Job satisfaction and autonomy at Wave IV are reported on ordered scales (of 0 to 4 and 0 to 3 respectively). A satisfaction score of 4 (extremely satisfied) is reported by 25.7% of men and 24.9% of women. The top autonomy score of 3 (always free to make important decisions at work) is declared by 37.4% of men and 29.8% of women. We pick up managerial responsibilities by a dummy for the respondent supervising one or more employees: this applies to 41.4% of men but only 31.7% of women. Men's annual earnings are over $9000 higher than women's, which difference is not explained by the number of hours worked per week.

These data allow us to see which types of sport affect which labour-market outcomes, and for whom (as we carry out separate analyses by gender). We here concentrate on the results for autonomy and managerial responsibilities, as those with respect to the other outcomes were more mixed. In particular, we find no significant relationship between

[6]Wave I respondents are asked "At the most recent grading period, what was your grade in each of the following subjects?". These grades were reported for English, Maths, History and Science, and are coded in a standard way, with 4 corresponding to an A, through to 1 for a D.

[7]English and History may require more time to master: the learning process in Maths and Science is more systematic.

[8]This "wage" includes wages or salary and tips, bonuses, and overtime pay, as well as the income from self-employment.

sport at school and the probability of being a worker (in line with the findings in Lechner (2009) using German data) or job satisfaction. Most of the estimated sport coefficients in a wage equation were also not significant. The finding that sport at Wave I does not predict employment at Wave IV means that we can analyse the characteristics of adult jobs without worrying about a selection bias due to childhood sport.

Completed education at Wave IV (almost all respondents have finished their education by this time) is measured by seven dummy variables for: Less than high school; High school; Training; College; Master; PhD; and Professional school.[9] Women in Add Health Wave IV are on average better-educated than are men, and had better grades at school in each of the disciplines measured at Wave I. Despite this academic advantage, girls were less happy at school than were boys (measured on a one to five scale, where 5 corresponds to "very happy"). Our regressions also include a number of other control variables, such as work experience, ethnicity, health (on an ordered scale), and age.[10]

A first simple exercise is to compare the characteristics of those who practised sport at school to those who did not. Mean-equality tests show that those with sport at school achieved better grades at school (Wave I) and end up significantly more educated (Wave IV); they are also significantly healthier. Crucially, they are also significantly more successful on the labour market 13 years later compared to those who did not practise sport. We now attempt to set out this relationship in a regression framework.

4 Sport and the Labour Market – A Causal Relationship?

Our aim here is to identify the effect of childhood sports participation on adult labour-market outcomes. The literature correctly notes that sports participation and labour-market success may well be driven by the same individual characteristics. One obvious ploy here is to control for the observables that we suspect might simultaneously determine childhood sports and labour-market success. The more difficult task relates to unobservables: here we appeal to information on siblings contained in the Add Health data in order to help estimate a more causal relationship.

4.1 Baseline Results

It is well-known that contemporaneous sport participation is correlated with both income and education. There is a problem of reverse causality, whereby those who are richer now either don't have the time to engage in sport, or are better able to pay for

[9] "Professional schools" are for example Law or Medical schools.

[10] The description of all of the variables which we use here can be found in Table 1.

sport-club memberships, for example. We here avoid this problem by appealing to data on individual sport participation 13 years prior to the labour-market outcomes at Wave IV.

Sport also affects health, body shape and the returns to schooling. These are components of human capital, which helps to determine individual labour-market outcomes. We consider this indirect impact of sport on the labour market by separately including controls for Wave IV education, health and household composition. As this is part of the effect we want to evaluate, we also consider results without these variables in the regression to have a rough estimate of these channels of influence.

We run separate regressions for men and women. We do so because we do not know if the effect of sport on non-cognitive skills is the same by gender. The descriptive statistics in Table 2 certainly suggest considerable differences in both sport participation at school and adult labour-market outcomes. We estimate the following equation:

$$Job\ Characteristics_{i,W4} = f(sport_{i,W1}, X_{i,W1}, Z_{i,W4}) \tag{1}$$

where $X_{i,W1}$ is a vector of control variables measured at Wave I (including age, ethnicity, and the education level of the most-educated parent) and $Z_{i,W4}$ a vector of Wave-IV mediators (including education, work experience, number of children, health and number of working hours). We separately estimate the probability of having full autonomy at work (defined as the maximum score of 3 on the 0 to 3 scale) and of being a manager via logit regressions.

As a number of different students are interviewed within each school in the Wave I Add Health data, we are able to introduce school fixed effects. This helps address some of the endogeneity concerns. In particular, it can be imagined that children who live in richer areas go to schools with better-quality infrastructure, where more sport can be practised. It is not surprising that children from richer backgrounds obtain better jobs, which then introduces a standard omitted-variable bias into our estimation. The use of school fixed effects allows us to compare the different adult labour-market outcomes of children who went to the same school.

4.2 Channels

4.2.1 Mediating Variables

We identify a number of mediating variables, whose influence is tested by their successive introduction and removal from the baseline specification. Sport may affect labour-market outcomes via: *i*) health; *ii*) cognitive and non-cognitive skills; and *iii*) the enlargement

and diversification of the individual's network. In the baseline estimation we control for *i*) and *ii*), so that any residual sport effect operates other than via education and health. It is also possible that sport affect body shape, and therefore success on the marriage market. Education is equally correlated with years of work experience. We hence successively add and remove each of these variables from the baseline specification to see how they affect the estimated coefficient on Wave I sport participation.

4.2.2 Additional Covariates

The Add Health data contain a wide variety of questions regarding the child's feelings in Wave I. We do have an issue with missing values, as many of these are not answered, which poses problems in terms of both potential bias and sample size. As such, these variables are not included in the baseline specification, but are introduced separately later. There are three groups of additional variables. First, the child's feelings towards school (whether they are happy there) and self-confidence. Levy-Garboua *et al.* (2006) show that adolescents who are unhappy at school also engage in risky behaviours which negatively affect their subsequent labour-market outcomes. We consider the following perception variables: "I have a lot of energy", "I have a lot of good qualities", "I am physically fit", "I have a lot to be proud of", "I like myself just the way I am", "I am happy to be at this school" and "I feel like I am doing everything just right". The second set of variables refer to the child's ability, as reflected in their self-declared grades in Maths, Science, English and History. Last, we have a measure of the child's popularity. At Wave I, the children declare both up to three friends and their best friend. We check whether the friends identified at school also declare the respondent as one of their friends (so that they are reciprocal friends); a similar strategy is adopted for the best friend.

A last additional set of covariates refer to other extra-curricular activities which are in some sense analogous to sport (meeting friends, hobbies, and watching TV). The introduction of these new variables may help us to understand how sport affects labour-market outcomes. For example, if the predominant channel is that of networking, then it is reasonable to imagine a similar labour-market effect from other activities, such as the frequency of meeting friends.

4.3 Endogeneity

Endogeneity is an issue in almost all econometric estimation using survey data. We are able to eliminate one source of endogeneity by using information on sports participation 13 years before the labour-market outcomes. We can also introduce school fixed effects, which will pick up any unobservables at the local-area or school level. There still remains the potential endogeneity issue that those who practise sport more have some unobservable individual characteristics (such as non-cognitive skills)

which also yield labour-market success. These introduce a standard positive omitted-variable bias: childhood sport and adult labour-market success will be correlated, but not necessarily causally so. This distinction is of course key for policy: encouraging sports at school to help ensure adult success will only work if the former causes the latter.

Our approach to this final source of endogeneity is to use the structure of the Add Health data to compare sibling outcomes, so that family environment is held constant. This method was developed and applied by Ashenfelter & Zimmerman (1997), Bronars & Oettinger (2006), Kosteas (2011) and more recently Lundborg *et al.* (2013). By doing so, we capture any unobservable characteristics, which are developed at the family level. The sibling regressions (which do not include race or school fixed effects, for obvious reasons) are estimated using the following specification:

$$\Delta Job\ Characteristics_{i,W4} = f(\Delta sport_{i,W1}, \Delta X^*_{i,W1}, \Delta Z^*_{i,W4}) \tag{2}$$

Here the $X^*_{i,W1}$ and $Z^*_{i,W4}$ vectors are similar in nature to those used in equation (1) above. Here they refer to differences between siblings, so race is not included, for example, and age is picked up by a dummy variable indicating the older child in the pair. Black *et al.* (2005a) show that birth order is significantly and negatively correlated with child education, independently of the size of the family.

A substantial part of unobservables are likely related to the child's socio-economic environment and parents. The notion of social reproduction in Bourdieu (1979) refers to the familial environment (within a social class) including tastes, parenting style and many socio-economic variables. Work on the effect of parental socio-economic background on child outcomes emphasises permanent income (Blau, 1999) and household (especially maternal) characteristics. Black *et al.* (2005b) use Norwegian data to argue that the correlation between parental and child education is not causal (except between mothers and sons). In their opinion the selection of individuals with certain characteristics into education is responsible for this correlation. Our family fixed effect controls for the part of the selection effect due to shared socio-economic background (including wealth, human capital and social networks). Parental sport itself will also matter: child sport participation is strongly affected by that of the parents. Raudsepp (2006) finds that fathers act as role model for sons but not daughters. Our separate estimation by gender (within sibling pairs) should shed some light on this type of effect.

Add Health also provides data on twins. It is tempting to argue that this is the ultimate way of solving endogeneity due to unobservables. However, Bound & Solon (1999) argue that there is no reason why twins should necessarily be preferred to normal siblings for the identification of causal relationships. In order for an effect to be estimated,

the twins have to differ in the treatment, and this difference is most likely endogeneous. We do estimate results within twin pairs, who share both genes and parents (although there turn out to be only relatively few of these). In general, the comparison of the adult outcomes of children from the same household allows us to control for a substantial amount of endogeneity.

5 Results

5.1 Manager

5.1.1 All

The baseline results by gender and school type appear in Table 3: these show the marginal effects from logit regressions on being a manager in Add Health Wave IV. For men, one more individual sport episode per week in middle school is associated with a 1.6 percentage-point higher probability of being a manager 13 years later; the analogous figure for team sports in high school is similar at 1.4 percentage points. For women, individual sports at high school increase the probability of being a manager 13 years later. These are arguably fairly large effects. The estimated effect of moving from the lowest (no) to the highest weekly sport category involves a cardinal change in sport intensity of 5.5, and thus an estimated marginal effect of nine percentage points on the probability of being a manager. The mediation analysis of this relationship reveals very little evidence of sport being channeled via health or household composition, with some suggestion of an effect (although only slight in size) via working experience and education (results available on request). The other right-hand side variables in Table 3 show that Black respondents have a consistently lower probability of occupying managerial positions, while education and work experience attract positive coefficients in this respect. Health (as measured at Wave IV) only seems to affect women's managerial chances.

Table 4 then asks whether sport acts differently from other extra-curricular activities (meeting friends, hobbies, and watching TV: each measured as weekly intensity at Wave I). In each panel we introduce these separately, and then finally all together. There is overall no evidence that these other activities change the estimated relationship between childhood sport and adult managerial position. When all activities are entered together, only two non-sport activities attract significant estimated coefficients (positive for boys' meeting friends in middle school, and negative for girls' TV watching in middle school).

Finally, we can also add the additional covariates described in Section 4.2.2. As intimated there, the addition of these behaviour and ability variables from Wave 1 does sharply reduce the sample size and potentially selects children who are more confident (and therefore answered the personal questions). We compare these new estimates

with those from the baseline specification but run on the same reduced sample. The results (available on request) underline that some of these new covariates do indeed predict managerial positions at Wave IV; however, their inclusion does not affect the previously-noted significance of the sport variables in Table 3.

Our results are identified by comparing the labour-market outcomes of different individuals. As we have school fixed effects, these comparisons concern individuals who went to the same school. Even so, they mostly come from different families, so that there is the continued possibility of unobserved family or parental characteristics which drive both childhood sport and adult labour-market success. We therefore now turn to the analysis of siblings in the Add Health data.

5.1.2 Siblings

Sibling analysis is not carried out separately for school type (middle and high school) for sample-size reasons; school type is entered as a control variable. We do separately consider various types of sibling pairs, from the widest – biological – definition to the narrowest (identical twins).

Table 5 shows the results for both genders. The dependent variable here is the difference in being a manager between siblings, and so takes on the values -1, 0 and 1. We estimate this equation by OLS, so that the estimated coefficients in Table 5 are the marginal effects. The results show that the difference in the frequency of individual sports between siblings when at school is positively and significantly correlated with the probability of being a manager 13 years later, with an estimated coefficient that is quite stable across columns (and similar in terms of marginal effect to the significant estimated coefficients in Table 3). We only lose significance here in the final column, where we consider the (fewer than 200 observations on) identical twins in Add Health data. There is no significant impact in Table 5 for team sports. With respect to the other control variables, we continue to find the same correlations with education, hours of work and work experience as in Table 3.

Table 6 then splits the analysis by gender, comparing successively related boys, related girls, and then related boys and girls. The estimated coefficients are systematically larger and more significant between related boys. In particular, there is no evidence that childhood sport differences between related girls predict adult labour-market success here, although we do find some evidence along these lines for the comparison between related boys and girls in the bottom panel of the table. As in Table 5, the significant differences here are all found for individual sport when at school.

As in the main sample, controlling for other leisure activities at Wave I does not affect

the significance of the estimates on our sibling sport variables. One consistent result here is that the difference in the time spent on hobbies between siblings at Wave I is positively and significantly linked to their probability of holding a managerial position at Wave IV (again, except for identical twins). Last, we find no evidence here of any mediation via health, education, working experience or household composition.

5.2 Autonomy

5.2.1 All

Table 7 shows the marginal effects from the logit estimation of having the highest level of autonomy at work in Wave IV (which is declared by declared by 37% of men and 30% of women). Childhood sport predicts adult labour-market autonomy, but only significantly so for men. In particular, team sports matter much more for individual sport for boys. In both cases the correlation is the strongest when the sport was practised at high school (as opposed to middle school). The estimated coefficients on the other control variables to a large extent mirror those for being a manager in Table 3.

Controlling for other activities while at school in Table 8 again does not affect the relationship between childhood sport and adult job autonomy: there does seem to be something specific about practising sport. The results show a strong correlation between adult autonomy and childhood time spent by girls on hobbies; for boys this correlation is with childhood time spent meeting friends (positive) and watching TV (negative). The mediation analysis suggests that a small part of the correlation between sport and autonomy transits via health, for both men and women.

5.2.2 Siblings

The results of the sibling estimation for adult autonomy for both genders appear in Table 9. There are no significant correlations with respect to individual sports here. On the contrary, the estimated coefficients on sibling differences in team sports are significant until we hit the small twin samples, and with an estimated coefficient that is notably stable across the five different columns. As in Table 7, hours of work, education and work experience are significantly correlated with autonomy. Table 9 also shows that women in Wave IV of the Add Health data are significantly less likely to report having full autonomy at work. Table 10 separates these estimations by sibling gender. All of the team sport effect found in Table 9 comes from the differences between related boys. The analysis of other activities between siblings at Wave I (as for managerial responsibilities) again underlines the importance of time spent on hobbies in childhood. Last, the mediation analysis between siblings produced no significant correlations.

6 Conclusion

This paper has used long-run American panel data to show that childhood sporting activities are often correlated with the level of autonomy and managerial responsibilities reported by the same individuals at work 13 years later. Different types of childhood sport affect men's and women's adult labour-market outcomes differently. Broadly speaking, team sport plays a larger role for boys, and individual sport for girls. One reading is that firms do not value or expect the same skills for men and women. In particular, sport seems to be valued to the extent that it reflects observed differences in behaviour (girls are slightly more likely to participate in individual sports; boys are much more likely to engage in team sports).

We appeal to a variety of strategies to address the endogeneity of childhood sport, via school fixed effects, controls for childhood behaviour, happiness and popularity, and within-sibling estimation. We also control for a variety of other activities undertaken by children when interviewed at Wave I, with the idea that unobservable differences between individuals (and between siblings) may well be picked up in time spent socialising, on hobbies, and watching TV. We largely continue to find significant correlations between childhood sport and labour-market outcomes 13 years later. In particular, the relation between sport and managerial responsibilities is robust to both school and family fixed effects. With respect to our two labour-market outcome variables, it is notable that individual sport is more strongly correlated with future managerial responsibilities, while autonomy is more strongly predicted by childhood participation in team sports.

We cannot make strong claims as to the channels via which this effect works: our mediation analysis reveals mostly insignificant effects from health, education, work experience and number of children. It is possible that sport affect labour-market outcomes via networking and signalling (where sport sends out signals about non-cognitive skills), neither of which we can test explicitly. It is worth noting however that the frequency of meeting friends in Wave I, which also involve the creation of networks, has either a smaller effect or no effect at all on adult labour-market outcomes. Either the type of networks created by sport are inherently more useful in the labour-market sense than those that are created by the other activities that schoolchildren undertake, or sport does indeed carry a strong signal about the type of person who engages in it, which is at least consistent with its prominent appearance in the CVs of many young labour-market entrants.

References

Anderson, Deborah J. 2001. "If you let me play": the effects of participation in high school athletics on students' educational and analyse market success.

Ashenfelter, Orley, & Zimmerman, David. 1997. Estimates of the returns to schooling from sibling data: fathers, sons and brothers. *Review of Economics and Statistics*, **79**(1), 1–9.

Barron, John M., Ewing, Bradley T., & Waddell, Glen R. 2000. The Effects Of High School Athletic Participation On Education And Labor Market Outcomes. *Review of Economics and Statistics*, **82**(3), 409–421.

Black, Sandra E., Devereux, Paul J., & Salvanes, Kjell G. 2005a. The More the Merrier? The Effect of Family Size and Birth Order on Children's Education. *Quarterly Journal of Economics*, **120**(2), 669–700.

Black, Sandra E., Devereux, Paul J., & Salvanes, Kjell G. 2005b. Why the Apple Doesn't Fall Far: Understanding Intergenerational Transmission of Human Capital. *American Economic Review*, **95**(1), 437–449.

Blau, David M. 1999. The Effect Of Income On Child Development. *Review of Economics and Statistics*, **81**(2), 261–276.

Bound, John, & Solon, Gary. 1999. Double trouble: on the value of twins-based estimation of the return to schooling. *Economics of Education Review*, **18**(2), 169–182.

Bourdieu, Pierre. 1979. *La Distinction. Critique sociale du jugement.* Les Éditions de Minuit.

Breuer, Christoph. 2006. *Sportpartizicaption in Deutschland - ein demo-ökonomisches Modell.* Discussion Paper. DIW Berlin.

Breuer, Christoph, & Wicker, Pamela. 2008. Demographic and economic factors influencing inclusion in the German sport system - a microanalysis of the years 1985 to 2005. *European Journal for Sport and Society*, **5**(1), 35–43.

Bronars, Stephen G., & Oettinger, Gerald S. 2006. Estimates of the return to schooling and ability: evidence from sibling data. *Labour Economics*, **13**(1), 19–34.

Cobb-Clark, Deborah A., & Schurer, Stefanie. 2012. The stability of big-five personality traits. *Economics Letters*, **115**(1), 11–15.

Downward, Paul. 2007. Exploring the Economic Choice to Participate in Sport: Results from the 2002 General Household Survey. *International Review of Applied Economics*, **21**(5), 633–653.

Ewing, Bradley T. 1998. Athletes and work. *Economics Letters*, **59**(1), 113–117.

Farrell, Lisa, & Shields, Michael A. 2002. Investigating the economic and demographic determinants of sporting participation in England. *Journal of the Royal Statistical Society Series A*, **165**(2), 335–348.

Felfe, Christina, Lechner, Michael, & Steinmayr, Andreas. 2011. Sports and Child Development. *CEPR Discussion Paper No. 8523*.

Gustafson, Sabrina L., & Rhodes, Ryan E. 2006. Parental Correlates of Physical Activity in Children and Early Adolescents. *Sports Medicine*, **36**(1), 79–97.

Hamermesh, Daniel S, & Biddle, Jeff E. 1994. Beauty and the Labor Market. *American Economic Review*, **84**(5), 1174–94.

Heckman, James J., & Kautz, Tim. 2012. Hard evidence on soft skills. *Labour Economics*, **19**(4), 451–464.

Hovemann, Gregor, & Wicker, Pamela. 2009. Determinants of sport participation in the European Union. *European Journal for Sport and Society*, **6**(1), 49–57.

Humphreys, Brad, & Ruseski, Jane. 2011. An Economic Analysis of Participation and Time Spent in Physical Activity. *The B.E. Journal of Economic Analysis & Policy*, **11**(1).

Kosteas, Vasilios D. 2011. High School Clubs Participation and Future Supervisory Status. *British Journal of Industrial Relations*, **49**(s1), 181–206.

Kosteas, Vasilios D. 2012. The Effect of Exercise on Earnings: Evidence from the NLSY. *Journal of Labor Research*, **33**(2), 225–250.

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

Levy-Garboua, Louis, Lohéac, Youenn, & Fayolle, Bertrand. 2006. Preference Formation, School Dissatisfaction and Risky Behavior of Adolescents. *Journal of Economic Psychology*, **27**(1), 165–183.

Long, James E., & Caudill, Steven B. 1991. The impact of Participation in Intercollegiate Athletics on Income and Graduation. *Review of Economics and Statistics*, **73**(3), 525–531.

Lundborg, Petter, Nilsson, Anton, & Rooth, Dan Olof. 2013. Adolescent Health and Adult Earnings: Evidence from a Large Sample of Siblings and Twins.

Mobius, Markus M., & Rosenblat, Tanya S. 2006. Why Beauty Matters. *American Economic Review*, **96**(1), 222–235.

Pfeifer, Christian, & Cornelißen, Thomas. 2010. The impact of participation in sports on educational attainment–New evidence from Germany. *Economics of Education Review*, **29**(1), 94–103.

Raudsepp, Lennart. 2006. The Relationship between socio-economic status, parental support and adolescent physical activity. *Acta Paediatrica*, **95**(1), 93–98.

Rees, Daniel, & Sabia, Joseph. 2010. Sports participation and academic performance: Evidence from the National Study of Adolescent Health. *Economics of Education Review*, **29**(5), 751–759.

Rooth, Dan Olof. 2011. Work Out or Out of Work: The Labor Market Return to Physical Fitness and Leisure Sport Activities. *Labour Economics*, **18**(3), 399–409.

Spence, James. 1973. Job Market Signaling. *Quarterly Journal of Economics*, **87**(3), 355–374.

Stevenson, Betsey. 2010. Beyond the Classroom: Using Title IX to Measure the Return to High School Sports. *Review of Economics and Statistics*, **92**(2), 284–301.

Table 1: Summary statistics: Working (at least 10 hours per week) versus Not Working

	Not Working			Working					Diff	
	Obs	**Mean**	**Std. Dev.**	**Obs**	**Mean**	**Std. Dev.**	**Min**	**Max**	**Size**	**Signif.**
Activities (All no. of times per week at W1)										
Individual sport	5651	2.85	1.97	10549	2.88	1.99	0	5.5	0.03	
Team sport	5651	2.32	2.10	10550	2.45	2.12	0	5.5	0.13	***
Meeting friends	5651	3.46	1.94	10549	3.53	1.91	0	5.5	0.07	**
Hobbies	5651	2.52	2.01	10550	2.56	1.97	0	5.5	0.04	
Watching TV	5651	4.21	1.70	10550	4.22	1.71	0	5.5	0.01	
Personal characteristics										
Race: Black	5564	0.18	0.38	10554	0.18	0.38	0	1	0.00	
Race: Asian	5564	0.06	0.23	10554	0.05	0.23	0	1	0.00	
Race: Indian	5564	0.03	0.16	10554	0.02	0.15	0	1	-0.01	**
Race: White	5564	0.51	0.50	10554	0.54	0.50	0	1	0.03	***
Race: Mixed	5564	0.23	0.42	10554	0.21	0.40	0	1	0.02	***
Male	5654	0.41	0.49	10554	0.49	0.50	0	1	0.09	***
Health status at W4	5654	2.37	0.97	10554	2.34	0.89	1	5	-0.03	*
Age at W4	5654	28.45	1.81	10554	28.56	1.78	24	34	0.11	***
No. of children at W4	5653	1.08	1.24	10554	0.88	1.11	0	8/7	-0.20	***
Years of work experience at W4	5134	8.23	3.43	10158	8.93	3.06	0	24/22	0.70	***
Ed: Less than high school at W4	5661	0.10	0.30	10554	0.07	0.25	0	1	-0.04	***
Ed: High school at W4	5661	0.56	0.50	10554	0.54	0.50	0	1	-0.01	
Ed: Training at W4	5661	0.06	0.23	10554	0.07	0.25	0	1	0.01	***
Ed: College at W4	5661	0.21	0.41	10554	0.24	0.43	0	1	0.03	***
Ed: Master at W4	5661	0.06	0.23	10554	0.06	0.24	0	1	0.00	
Ed: PhD at W4	5661	0.01	0.08	10554	0.01	0.09	0	1	0.00	
Ed: Professional school at W4	5661	0.01	0.10	10554	0.01	0.11	0	1	0.00	
Education of the most educated parent	4887	5.88	2.41	9098	6.09	2.29	0	9	0.21	***
In high school at W1 (versus middle school)	5493	0.52	0.50	10317	0.55	0.50	0	1	0.04	***
Additional covariates (All at W1)										
Have a lot of energy (1-5)	3792	2.05	0.93	7224	2.02	0.91	1	5	-0.03	
Have a lot of good qualities (1-5)	3725	1.87	0.87	7097	1.83	0.84	1	5	-0.03	*
Physically fit (1-5)	3717	2.20	1.02	7085	2.14	1.01	1	5	-0.06	*
Have a lot to be proud of (1-5)	3712	1.89	0.94	7078	1.87	0.94	1	5	-0.01	
Like myself just the way I am (1-5)	3716	2.16	1.12	7069	2.19	1.13	1	5	0.03	
Feel like I am doing everything just right (1-5)	3700	2.70	1.04	7063	2.72	1.05	1	5	0.02	
Feel loved and wanted (1-5	3702	2.04	1.00	7064	2.06	1.00	1	5	0.02	
Grade in English (A-D	5385	2.85	0.96	10137	2.81	0.96	1	4	0.04	***
Grade in Maths (A-D)	5152	2.68	1.05	9679	2.64	1.04	1	4	0.04	**
Grade in History (A-D)	4895	2.83	1.01	9095	2.87	1.00	1	4	-0.03	*
Grade in Science (A-D)	4885	2.83	1.01	9076	2.79	1.01	1	4	0.03	*
Best friend reciprocal	2623	0.40	0.49	4971	0.42	0.49	0	1	0.02	*
Friend reciprocal	2623	0.68	0.47	4971	0.70	0.46	0	1	0.02	*
Happy at school (1-5)	3733	3.58	1.21	7097	3.55	1.22	1	5	0.03	

Notes: * significant at 10%; ** significant at 5%; *** significant at 1%. When two maximum figures appear, the first refers to those not working and the second to workers.

Table 2: Summary statistics for outcomes and activities by gender

	Women			Men					Diff	
	Obs	**Mean**	**Std. Dev.**	**Obs**	**Mean**	**Std. Dev.**	**Min**	**Max**	**Size**	**Signif.**
Outcomes at W4										
Having paid work	8690	0.61	0.49	7518	0.69	0.46	0	1	0.08	***
Log of the annual wage	5078	9.83	1.90	5002	10.24	1.59	0	13.7/13.8	0.41	***
Being a manager	5330	0.32	0.47	5221	0.41	0.49	0	1	0.10	***
Full autonomy at work	5329	0.30	0.46	5219	0.37	0.48	0	1	0.07	***
Job satisfaction (0-4)	5328	2.91	0.90	5221	2.92	0.90	0	4	0.01	
Activities (all no. of times per week) at W1										
Individual sport	8688	2.89	1.92	7518	2.84	2.04	0	5.5	-0.05	
Team sport	8688	1.86	1.96	7519	3.03	2.12	0	5.5	1.17	***
Meeting friends	8688	3.42	1.96	7518	3.60	1.87	0	5.5	0.19	***
Hobbies	8688	2.42	1.94	7519	2.69	2.03	0	5.5	0.26	***
Watching TV	8688	4.14	1.73	7519	4.30	1.67	0	5.5	0.16	***

Notes: * significant at 10%; ** significant at 5%; *** significant at 1%. When two maximum figures appear, the first refers to women and the second to men.

Table 3: Manager Regressions

	Men			*Women*		
	All	**Middle sch.**	**High sch.**	**All**	**Middle sch.**	**High sch.**
Individual sport, times per week at W1	0.003	0.016**	-0.005	0.009**	0.002	0.016***
	(0.004)	(0.006)	(0.006)	(0.004)	(0.006)	(0.006)
Team sport, times per week at W1	0.010**	0.002	0.014**	-0.000	-0.003	0.004
	(0.004)	(0.006)	(0.006)	(0.004)	(0.006)	(0.006)
Hours worked per week	0.011***	0.009***	0.012***	0.011***	0.010***	0.011***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Race: Mixed	0.010	0.024	0.003	-0.028	-0.061*	-0.019
	(0.021)	(0.033)	(0.028)	(0.021)	(0.033)	(0.030)
Race: Black	-0.097***	-0.076*	-0.137***	-0.039	-0.075**	-0.033
	(0.030)	(0.044)	(0.043)	(0.025)	(0.038)	(0.035)
Race: Asian	0.015	0.006	0.045	-0.029	-0.185**	0.027
	(0.045)	(0.078)	(0.056)	(0.043)	(0.092)	(0.054)
Race: Indian	-0.071	-0.095	-0.086	0.025	0.070	-0.026
	(0.057)	(0.095)	(0.081)	(0.060)	(0.100)	(0.084)
Health: Very good at W4	-0.022	-0.028	-0.028	0.035	-0.034	0.104***
	(0.022)	(0.034)	(0.031)	(0.022)	(0.032)	(0.032)
Health: Good at W4	-0.027	-0.016	-0.035	0.038	-0.024	0.106***
	(0.023)	(0.036)	(0.033)	(0.023)	(0.033)	(0.034)
Health: Fair at W4	0.010	0.043	-0.015	0.042	-0.022	0.070
	(0.034)	(0.052)	(0.047)	(0.033)	(0.048)	(0.049)
Health: Poor at W4	-0.091	0.018	-0.195	-0.092	-0.153	-0.152
	(0.119)	(0.178)	(0.167)	(0.113)	(0.159)	(0.190)
Age at W4	-0.074	0.281	-0.466	0.098	0.219	0.903*
	(0.138)	(0.398)	(0.446)	(0.138)	(0.387)	(0.470)
Age-squared at W4	0.001	-0.006	0.008	-0.002	-0.004	-0.015*
	(0.002)	(0.007)	(0.007)	(0.002)	(0.007)	(0.008)
No. of children at W4	0.007	0.014	0.007	0.010	0.015	0.008
	(0.008)	(0.014)	(0.011)	(0.007)	(0.011)	(0.011)
Years of work experience at W4	0.017***	0.025***	0.010*	0.008**	0.004	0.011**
	(0.004)	(0.006)	(0.005)	(0.004)	(0.005)	(0.005)
Ed: Less than high school at W4	0.019	0.009	0.004	-0.050	-0.025	-0.058
	(0.030)	(0.042)	(0.051)	(0.040)	(0.053)	(0.075)
Ed: Training at W4	-0.083**	-0.063	-0.083*	-0.092***	-0.131***	-0.065
	(0.035)	(0.058)	(0.048)	(0.031)	(0.046)	(0.044)
Ed: College at W4	0.071***	0.078*	0.069**	0.017	-0.001	0.027
	(0.023)	(0.040)	(0.031)	(0.021)	(0.032)	(0.029)
Ed: Master at W4	0.059	0.089	0.057	-0.035	-0.093*	-0.005
	(0.044)	(0.070)	(0.057)	(0.032)	(0.055)	(0.043)
Ed: PhD at W4	0.132	-0.249	0.250	0.264***	0.478***	0.254***
	(0.137)	(0.232)	(0.192)	(0.077)	(0.149)	(0.096)
Ed: Professional school at W4	0.184**	-0.031	0.320***	0.062	-0.049	0.146*
	(0.086)	(0.204)	(0.105)	(0.066)	(0.115)	(0.086)
Constant	0.781	0.053	2.965	-2.103	-6.369	-14.129**
	(1.989)	(0.000)	(0.000)	(1.982)	(0.000)	(6.990)
Observations	4,395	1,955	2,332	4,326	1,960	2,243

Notes: Robust standard errors in parentheses; * significant at 10%; ** significant at 5%; *** significant at 1%. School fixed effects are included. The coefficients displayed are the marginal effects. Middle sch. (High sch.) refers to individuals who were in middle (high) school when interviewed at W1. Omitted categories: White (race), Excellent (health), High School (education) and no school (parental education). All regressions include nine parental-education dummies.

Table 4: Manager Regressions, including other activities

Men	Middle school					High school				
Individual sport, times per week at W1	0.016**				0.015**	-0.005				-0.005
	(0.006)				(0.006)	(0.006)				(0.006)
Team sport, times per week at W1	0.002				0.000	0.014**				0.014**
	(0.006)				(0.007)	(0.006)				(0.006)
Meeting friends, times per week at W1		0.014**			0.012*		0.007			0.005
		(0.007)			(0.007)		(0.006)			(0.006)
Hobbies, times per week at W1			0.007		0.005			0.002		0.001
			(0.006)		(0.006)			(0.006)		(0.006)
Watching TV, times per week at W1				0.001	0.000				0.005	0.003
				(0.008)	(0.009)				(0.007)	(0.007)
Observations	1,955	1,955	1,956	1,956	1,955	2,332	2,332	2,332	2,332	2,332

Women	Middle school					High school				
Individual sport, times per week at W1	0.002				0.000	0.016***				0.015***
	(0.006)				(0.006)	(0.006)				(0.006)
Team sport, times per week at W1	-0.003				-0.003	0.004				0.003
	(0.006)				(0.006)	(0.006)				(0.006)
Meeting friends, times per week at W1		0.005			0.006		0.008			0.006
		(0.006)			(0.006)		(0.006)			(0.006)
Hobbies, times per week at W1			0.003		0.002			0.010*		0.006
			(0.006)		(0.006)			(0.006)		(0.006)
Watching TV, times per week at W1				-0.015**	-0.015**				-0.001	-0.002
				(0.007)	(0.007)				(0.006)	(0.006)
Observations	1,960	1,960	1,960	1,960	1,960	2,243	2,243	2,243	2,243	2,243

Notes: Robust standard errors in parentheses, * significant at 10%; ** significant at 5%; *** significant at 1%. School fixed effects are included. The coefficients displayed are the marginal effects. Other Covariates: work hours, ethnicity, health, number of children, education, parental education, age, work experience.

Table 5: Manager: Basic sibling estimation

	Sib. and Cousins	Half and Full sib.	Full sib.	Twins	Id. Twins
Sibling diff. in individual sport at W1	0.016***	0.019***	0.022***	0.018*	0.010
	(0.005)	(0.006)	(0.007)	(0.011)	(0.018)
Sibling diff. in team sport at W1	0.006	0.002	0.001	-0.004	0.005
	(0.005)	(0.006)	(0.007)	(0.011)	(0.019)
Sibling diff. in weekly hours of work at W4	0.006***	0.006***	0.005***	0.006***	0.006*
	(0.001)	(0.001)	(0.001)	(0.002)	(0.003)
Sibling diff. in age at W4	-0.051	-0.146	-0.174		
	(0.199)	(0.208)	(0.239)		
Sibling diff. in age-squared at W4	0.001	0.002	0.002		
	(0.003)	(0.004)	(0.004)		
Sibling diff. in working experience at W4	0.021***	0.020***	0.020***	0.019**	0.024*
	(0.004)	(0.004)	(0.004)	(0.009)	(0.013)
Sibling diff. in health at W4	0.004	-0.002	-0.004	-0.024	-0.019
	(0.012)	(0.013)	(0.014)	(0.021)	(0.037)
Sibling diff. in Ed: Less than high school at W4	-0.060*	-0.069	-0.076	-0.148	0.228
	(0.034)	(0.043)	(0.051)	(0.098)	(0.168)
Sibling diff. in Ed: Training at W4	-0.075**	-0.052	-0.073	-0.083	-0.145
	(0.038)	(0.043)	(0.049)	(0.067)	(0.104)
Sibling diff. in Ed: College at W4	0.090***	0.078**	0.084**	0.078	0.097
	(0.030)	(0.036)	(0.037)	(0.063)	(0.141)
Sibling diff. in Ed: Master at W4	0.078	0.053	0.058	0.072	-0.126
	(0.052)	(0.056)	(0.059)	(0.100)	(0.184)
Sibling diff. in Ed: Professional school at W4	0.124	0.081	0.084	0.407**	0.420*
	(0.098)	(0.105)	(0.105)	(0.171)	(0.223)
Sibling diff. in Ed: PhD at W4	0.204**	0.143	0.142	0.169	0.227
	(0.079)	(0.099)	(0.100)	(0.124)	(0.173)
Is the older sibling	-0.044	-0.071	-0.088*		
	(0.037)	(0.045)	(0.048)		
Female	-0.055*	-0.064*	-0.043	-0.042	
	(0.032)	(0.035)	(0.038)	(0.057)	
Sibling diff. in the no. of children at W4	-0.005	-0.006	-0.003	-0.047***	0.008
	(0.009)	(0.010)	(0.011)	(0.018)	(0.038)
Constant	0.060**	0.066**	0.055*	0.057	0.059
	(0.030)	(0.032)	(0.033)	(0.042)	(0.038)
Observations	1,853	1,504	1,269	475	180
R-squared	0.054	0.050	0.049	0.079	0.074

Notes: Robust standard errors in parentheses; * significant at 10%; ** significant at 5%; *** significant at 1%.

Table 6: Manager: Sibling estimation by gender

Related boys	**Sib. and Cousins**	**Half and Full sib.**	**Full sib.**	**Twins**	**Id. Twins**
Sibling diff. in individual sport at W1	0.032***	0.034***	0.033***	0.037**	0.006
	(0.010)	(0.011)	(0.011)	(0.018)	(0.021)
Sibling diff. in team sport at W1	0.005	0.006	0.006	-0.017	0.001
	(0.012)	(0.014)	(0.015)	(0.022)	(0.030)
Observations	508	431	382	167	83
R-squared	0.082	0.080	0.087	0.156	0.125
Related girls	**Sib. and Cousins**	**Half and Full sib.**	**Full sib.**	**Twins**	**Id. Twins**
Sibling diff. in individual sport at W1	0.012	0.017	0.016	0.012	0.015
	(0.010)	(0.011)	(0.014)	(0.023)	(0.032)
Sibling diff. in team sport at W1	0.011	0.012	0.009	0.006	0.006
	(0.008)	(0.010)	(0.012)	(0.019)	(0.030)
Observations	629	499	421	176	97
R-squared	0.040	0.053	0.055	0.096	0.089
Related boys and girls	**Sib. and Cousins**	**Half and Full sib.**	**Full sib.**	**Twins**	
Sibling diff. in individual sport at W1	0.013	0.017*	0.023*	0.004	
	(0.008)	(0.010)	(0.012)	(0.020)	
Sibling diff. in team sport at W1	-0.004	-0.016*	-0.018*	-0.001	
	(0.009)	(0.009)	(0.010)	(0.019)	
Observations	716	574	466	132	
R-squared	0.088	0.085	0.081	0.079	

Notes: Robust standard errors in parentheses; * significant at 10%; ** significant at 5%; *** significant at 1%. Covariates: the differences in age, health, number of children, work experience, level and type of education and number of working hours, as well as gender and a dummy for "older" (when relevant).

Table 7: Full Autonomy Regressions

	Men			*Women*		
	All	**Middle sch.**	**High sch.**	**All**	**Middle sch.**	**High sch.**
Individual sport, times per week at W1	0.006	0.012*	0.003	0.005	0.003	0.008
	(0.004)	(0.006)	(0.006)	(0.004)	(0.006)	(0.005)
Team sport, times per week at W1	0.012***	0.017***	0.012**	0.005	0.003	0.006
	(0.004)	(0.006)	(0.005)	(0.004)	(0.006)	(0.006)
Hours worked per week	0.006***	0.006***	0.005***	0.004***	0.003***	0.005***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Race: Mixed	0.020	0.034	0.003	-0.007	-0.018	0.001
	(0.020)	(0.032)	(0.027)	(0.020)	(0.033)	(0.028)
Race: Black	0.009	0.064	-0.041	0.008	0.038	-0.011
	(0.028)	(0.042)	(0.041)	(0.024)	(0.037)	(0.033)
Race: Asian	0.015	0.039	-0.000	-0.028	-0.026	-0.064
	(0.043)	(0.078)	(0.055)	(0.044)	(0.075)	(0.055)
Race: Indian	-0.034	-0.161*	0.051	0.085	0.052	0.080
	(0.056)	(0.096)	(0.075)	(0.057)	(0.099)	(0.072)
Health: Very good at W4	-0.067***	-0.061*	-0.082***	-0.099***	-0.108***	-0.091***
	(0.021)	(0.032)	(0.029)	(0.021)	(0.031)	(0.028)
Health: Good at W4	-0.086***	-0.111***	-0.073**	-0.091***	-0.083**	-0.096***
	(0.022)	(0.034)	(0.030)	(0.021)	(0.032)	(0.030)
Health: Fair at W4	-0.055*	0.032	-0.130***	-0.105***	-0.110**	-0.121***
	(0.033)	(0.049)	(0.047)	(0.032)	(0.047)	(0.046)
Health: Poor at W4	-0.059	0.062	-0.323*	-0.040	0.008	-0.062
	(0.101)	(0.164)	(0.185)	(0.081)	(0.117)	(0.121)
Age at W4	0.001	0.455	-0.639	0.073	0.618	0.346
	(0.132)	(0.372)	(0.434)	(0.129)	(0.420)	(0.415)
Age-squared at W4	-0.000	-0.009	0.010	-0.001	-0.012	-0.006
	(0.002)	(0.007)	(0.007)	(0.002)	(0.008)	(0.007)
No. of children at W4	0.022***	0.035***	0.013	0.021***	0.017	0.028***
	(0.008)	(0.013)	(0.010)	(0.007)	(0.011)	(0.010)
Years of work experience at W4	0.011***	0.014**	0.009*	0.007*	0.007	0.008
	(0.004)	(0.006)	(0.005)	(0.004)	(0.005)	(0.005)
Ed: Less than high school at W4	0.000	-0.023	0.036	-0.050	-0.064	-0.055
	(0.030)	(0.043)	(0.050)	(0.038)	(0.052)	(0.073)
Ed: Training at W4	0.022	0.001	0.038	0.026	0.001	0.044
	(0.032)	(0.052)	(0.044)	(0.027)	(0.041)	(0.039)
Ed: College at W4	0.053**	0.064*	0.027	-0.010	-0.023	0.003
	(0.022)	(0.036)	(0.030)	(0.021)	(0.032)	(0.029)
Ed: Master at W4	0.008	0.119*	-0.109*	-0.024	-0.008	-0.012
	(0.042)	(0.072)	(0.057)	(0.032)	(0.053)	(0.042)
Ed: PhD at W4	0.010		0.200	0.246***	0.282**	0.254***
	(0.123)		(0.144)	(0.074)	(0.126)	(0.092)
Ed: Professional school at W4	0.154**	0.238	0.144*	-0.119*	-0.114	-0.097
	(0.075)	(0.164)	(0.086)	(0.071)	(0.135)	(0.086)
Constant	-0.605	-6.330	5.940	-1.186	-11.224	-5.327
	(1.894)	(5.038)	(0.000)	(1.849)	(0.000)	(6.161)
Observations	4,388	1,946	2,328	4,330	1,972	2,236

Notes: Robust standard errors in parentheses; * significant at 10%; ** significant at 5%; *** significant at 1%. School fixed effects are included. The coefficients displayed are the marginal effects. Middle sch. (High sch.) refers to individuals who were in middle (high) school when interviewed at W1. Omitted categories: White (race), Excellent (health), High School (education) and no school (parental education). All regressions include nine parental-education dummies.

Table 8: Full Autonomy Regressions, including other activities

Men	Middle school					High school				
Individual sport, times per week at W1	0.012*				0.011*	0.003				0.002
	(0.006)				(0.006)	(0.006)				(0.006)
Team sport, times per week at W1	0.017***				0.016**	0.012**				0.011**
	(0.006)				(0.006)	(0.005)				(0.006)
Meeting friends, times per week at W1		0.002			-0.002		0.016**			0.016**
		(0.006)			(0.007)		(0.006)			(0.006)
Hobbies, times per week at W1			0.015**		0.011*			0.010*		0.009
			(0.006)		(0.006)			(0.005)		(0.005)
Watching TV, times per week at W1				-0.002	-0.002				-0.015**	-0.019***
				(0.008)	(0.008)				(0.006)	(0.006)
Observations	1,946	1,946	1,947	1,947	1,946	2,328	2,328	2,328	2,328	2,328

Women	Middle school					High school				
Individual sport, times per week at W1	0.003				0.001	0.008				0.006
	(0.006)				(0.006)	(0.005)				(0.006)
Team sport, times per week at W1	0.003				0.003	0.006				0.004
	(0.006)				(0.006)	(0.006)				(0.006)
Meeting friends, times per week at W1		0.004			0.003		0.005			0.004
		(0.006)			(0.006)		(0.005)			(0.005)
Hobbies, times per week at W1			0.012**		0.012*			0.014***		0.012**
			(0.006)		(0.006)			(0.005)		(0.005)
Watching TV, times per week at W1				-0.004	-0.005				-0.004	-0.004
				(0.007)	(0.007)				(0.006)	(0.006)
Observations	1,972	1,972	1,972	1,972	1,972	2,236	2,236	2,236	2,236	2,236

Notes: Robust standard errors in parentheses, * significant at 10%; ** significant at 5%; *** significant at 1%. School fixed effects are included. The coefficients displayed are the marginal effects. Other Covariates: work hours, ethnicity, health, number of children, education, parental education, age, work experience.

Table 9: Full Autonomy: Sibling estimation

	Sib. and Cousins	Half and Full sib.	Full sib.	Twins	Id. Twins
Sibling diff. in individual sport at W1	0.004	-0.001	-0.000	-0.000	-0.015
	(0.006)	(0.006)	(0.007)	(0.011)	(0.018)
Sibling diff. in team sport at W1	0.016***	0.018***	0.015**	0.011	0.019
	(0.005)	(0.006)	(0.006)	(0.013)	(0.022)
Sibling diff. in weekly hours of work at W4	0.003***	0.003**	0.003**	0.005**	0.003
	(0.001)	(0.001)	(0.001)	(0.002)	(0.004)
Sibling diff. in age at W4	-0.262	-0.272	-0.379*		
	(0.234)	(0.214)	(0.215)		
Sibling diff. in age-squared at W4	0.005	0.005	0.007*		
	(0.004)	(0.004)	(0.004)		
Sibling diff. in working experience at W4	0.012***	0.012**	0.012**	0.019**	0.010
	(0.004)	(0.005)	(0.005)	(0.008)	(0.014)
Sibling diff. in health at W4	-0.002	-0.006	-0.010	-0.020	-0.040
	(0.012)	(0.013)	(0.014)	(0.024)	(0.035)
Sibling diff. in Ed: Less than high school at W4	-0.003	-0.004	-0.029	-0.007	0.114
	(0.032)	(0.036)	(0.038)	(0.058)	(0.134)
Sibling diff. in Ed: Training at W4	0.037	0.077	0.076	0.083	-0.028
	(0.045)	(0.055)	(0.058)	(0.069)	(0.111)
Sibling diff. in Ed: College at W4	0.039	0.029	0.029	0.017	0.034
	(0.031)	(0.034)	(0.036)	(0.064)	(0.108)
Sibling diff. in Ed: Master at W4	0.028	0.007	0.034	0.172	0.125
	(0.061)	(0.064)	(0.065)	(0.111)	(0.168)
Sibling diff. in Ed: Professional school at W4	-0.002	-0.019	-0.014	-0.172	-0.269
	(0.094)	(0.106)	(0.107)	(0.179)	(0.260)
Sibling diff. in Ed: PhD at W4	0.332***	0.292**	0.295**	0.331*	0.264
	(0.106)	(0.119)	(0.119)	(0.187)	(0.225)
Is the older sibling	0.036	0.071	0.053		
	(0.040)	(0.051)	(0.056)		
Female	-0.058**	-0.049	-0.026	-0.060	
	(0.029)	(0.031)	(0.033)	(0.055)	
Sibling diff. in the no. of children at W4	0.013	0.010	0.010	0.017	0.029
	(0.009)	(0.009)	(0.011)	(0.019)	(0.027)
Constant	0.028	0.023	0.011	0.023	-0.012
	(0.029)	(0.029)	(0.030)	(0.037)	(0.041)
Observations	1,852	1,503	1,268	475	180
R-squared	0.032	0.032	0.028	0.055	0.054

Notes: Robust standard errors in parentheses; * significant at 10%; ** significant at 5%; *** significant at 1%.

Table 10: Full Autonomy: Sibling estimation by gender

Related boys	**Sib. and Cousins**	**Half and Full sib.**	**Full sib.**	**Twins**	**Id. Twins**
Sibling diff. in individual sport at W1	-0.000	-0.005	-0.007	-0.011	-0.027
	(0.009)	(0.011)	(0.012)	(0.017)	(0.024)
Sibling diff. in team sport at W1	0.025**	0.035***	0.034***	0.037	0.046
	(0.010)	(0.011)	(0.012)	(0.024)	(0.040)
Observations	507	430	381	167	83
R-squared	0.049	0.065	0.060	0.138	0.154
Related girls	**Sib. and Cousins**	**Half and Full sib.**	**Full sib.**	**Twins**	**Id. Twins**
Sibling diff. in individual sport at W1	0.013	0.010	0.010	0.008	-0.004
	(0.010)	(0.009)	(0.011)	(0.016)	(0.025)
Sibling diff. in team sport at W1	0.007	0.002	-0.000	0.002	-0.011
	(0.010)	(0.011)	(0.011)	(0.021)	(0.021)
Observations	629	499	421	176	97
R-squared	0.041	0.041	0.059	0.087	0.120
Related boys and girls	**Sib. and Cousins**	**Half and Full sib.**	**Full sib.**	**Twins**	
Sibling diff. in individual sport at W1	0.002	-0.005	-0.003	-0.007	
	(0.008)	(0.009)	(0.011)	(0.019)	
Sibling diff. in team sport at W1	0.010	0.007	0.001	0.003	
	(0.008)	(0.009)	(0.009)	(0.021)	
Observations	716	574	466	132	
R-squared	0.083	0.085	0.078	0.139	

Notes: Robust standard errors in parentheses; * significant at 10%; ** significant at 5%; *** significant at 1%. Covariates: the differences in age, health, number of children, work experience, level and type of education and number of working hours, as well as gender and a dummy for "older" (when relevant).