



IN SPORT, WE TRUST

Full Technical Report



because
sport works

Contents:

1. Executive Summary and Key Results	2
2. Methodology	8
3. Results in Full	17
4. Conclusion	34

Report and analysis undertaken by Jump Projects (2019)

1. Executive summary and key results

As the literature review from Leeds Beckett University accompanying this report demonstrates, this report builds on decades of work looking at the social benefits of collective endeavour and sport. From Robert Putnam's seminal book 'Bowling Alone' in 2000 to the 2016 Casey Review of opportunity and integration in the UK, we are understanding more and more about the role sport plays in building the bonds of trust that keep our communities healthy.

The last decade has also seen the addition of wellbeing and social cohesion measures into UK Government data and these offer the potential for new and innovative evidence and insight. Sported want to contribute to the evidence on the impact of community sport groups on young people, community development and social cohesion. Sport England's 2016 Evidence review cited this as "one of the hardest outcomes to evidence" and that "there is considerable scope for building the evidence base further around community development".

The essential components of this work are new, innovative analysis of large national population datasets in the UK to establish a robust link between the activity (sport group engagement) and social outcomes (community cohesion measures), allowing us to report new findings and identify gaps in the evidence. This analysis includes Sport England's own Active Lives dataset, which is analysed alongside other important datasets like Understanding Society, Taking Part and Community Life. All these datasets track participation in sports groups and almost all of the outcomes relevant to the DCMS Sporting Future Strategy and Sport England Evaluation Framework (health, wellbeing, individual development, community development). The target group of this study are young people aged 25 and below.

The importance of controlling for socio-economic factors in sport

Throughout the descriptive statistics in this report, sports participation and sport group membership is more common for higher socio-economic groups in society. It is important to acknowledge that affluence and earnings are positively associated with all of the five wellbeing outcomes from the 2016 Sporting Future strategy: physical health, mental wellbeing, individual development (education, skills, confidence), community development (trust, social mixing and volunteering) and economic development (e.g. monetary spend on playing, watching sport). Holding other things constant, on average, the more money you have, the more likely you are to be healthy, happy, trusting and interacting with a more diverse range of people and experiences.

Therefore, the crucial question for this research becomes: is it playing sport or membership of sports clubs that is improving health, wellbeing and trust, or is it the higher income, education or socio-economic background of sport players?

For this reason it is clearly spelled out in 2014 work for DCMS quantifying the impact of sports participation that:¹

Essential to this process is the ability to control for as many of the determinants of a given outcome as possible using regression analysis. It is the optimal method given the nature of the data.

This method of controlling for as many of the other factors that could cause the improvements is known as ‘multivariate regression analysis’. The 2014 DCMS paper was analysing just one dataset, Taking Part. For the present report, Jump have conducted exactly this form of multivariate analysis on all of the major UK datasets mentioned above. And this is vital in looking at the impact of sport for two reasons:

1. As above and as found throughout this report, sport and sports group membership is more common for higher socio-economic groups in society;
2. There is a strong indication that sport’s potential is to have a greater impact on lower socio-economic groups².

Community development measures

There are large datasets in the UK that ask a number of policy relevant measures for social and community cohesion outcomes. The What Works in Wellbeing Centre points to a number of measures in their useful impact and evaluation toolkit³ and these mirror the Jump approach in drawing on existing, validated datasets. The outcome measures analysed in this report have the following characteristics in common:

- **Self-reported engagement** in youth, sport, civic participation and volunteering linked to social outcomes (e.g. trust in others, social mixing and wellbeing).
- Cross-checked against **existing pre-validated data sources** (UK Government and NGO) on social outcomes, sport engagement, and youth impact.
- **Normative measures across wider policy** and practice landscape in sport, youth and volunteering.

The deficit of health, wellbeing and community development measures between high and lower socio-economic groups (SEG)

There is emerging evidence pointing to the fact that young people from lower SEG (defined as working in low skilled jobs) tend to be consistently less trusting of their neighbours (although they talk to them more) and are less socially connected with a less diverse pool of friends. Furthermore, they are considerably less likely to volunteer.

For example, in the Jump report “A Bit Rich: Why is volunteering biased towards higher SEG?”⁴, analysis of the Community Life data shows that lower SEG in the national sample have significantly lower odds of having mixed with people from different ethnic groups at least

¹

https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/304899/Quantifying_and_valuing_the_wellbeing_impacts_of_sport_and_culture.pdf

² Jump analysis of Understanding Society and Taking Part 2018 - to be published in 2019

³ <https://measure.whatworkswellbeing.org/homepage/recommended-questions/>

⁴ <https://jump-projects.com/our-work/>

once in the last 12 months (25% of the sample, compared to 37% of those from higher SEG) and also report slightly lower levels of trust in people in the local neighbourhood (40% agree that they trust many people, compared to the national average of 45%, and 51% with higher SEG). This is detailed in Table 4A (see below). There are also some interesting divergences in the trust that different groups hold for government institutions at the local and national level. This finding from Community Life is also consistently found in the other datasets we analysed and reported in tables 4A to 4D on pages 17-18.

Copy of Table 4A. Trust among lower socio-economic groups: Community Life (all ages)

	Lower SEG Odds ratio	Lower SEG (%, n/N)	Higher SEG (%, n/N)	Total observations (%, n/N)
Trust local council: a lot or a fair amount	1.070	63.5% (2221/3499)	61.1% (1590/2602)	62.5% (4181/6692)
Trust Parliament: a lot or a fair amount	0.953	31.4% (1107/3526)	32.9% (871/2647)	31.7% (2147/6773)
Trust police: a lot or a fair amount	0.867*	82.6% (2975/3602)	86.7% (2305/2658)	84.1% (5789/6882)
Trust many of the people in neighbourhood	0.771***	40.1% (3074/7668)	50.5% (3700/7322)	44.7% (7496/16754)

*Notes: Community Life 2012-2016, England. Lower socio-economic group defined as below sample median household income. Column 2: Logistic regression model. Heteroskedasticity-robust standard errors used. *** <1% significance; ** <5% significance; * <10% significance. Standard socio-demographic controls including household income, employment status, marital status, region etc. included but not reported in this table. Weights not applied due to sample size issues.*

The role of sport in addressing the deficit in trust, wellbeing and community outcomes: key results

In Table 1A below we see the outcomes of the regression analysis of sport group participation across all the relevant social/community and wellbeing outcomes for young people (aged 16-25) in sport. Consistently across different datasets we can see a **positive association between sports club membership and the following outcomes:**

- **Social connections** – having friends, number of close friends, relying on friends, satisfaction with friends
- **Trust** – trusting people in general, trusting neighbours (although we also note that there is no significant positive relationship for sport and trust in institutions)
- **Community cohesion** – talking to people in the local area, belonging to local area, satisfaction with local area
- **Volunteering**⁵ – consistently positive and significantly so across all datasets; in particular formal volunteering (Community Life)
- **Perceived ability to achieve goals** (perseverance) – in the Active Lives data.
- **Life satisfaction** (in all datasets) and happiness (wherever measured)
- **Health** (unsurprisingly) – in all datasets.

⁵ What we do not know from the data is whether they volunteer at the clubs where they are members, and whether they volunteered first and then became members or vice versa. This is impossible within the current datasets, but may be an area for future primary research.

Table 1A. Key associations between sports club membership and social outcomes

Outcome variable	OLS regression coefficient	Dataset	Interpretation
Number of close friends	0.345***	Understanding Society	Membership of a sport club / organisation is associated with 0.345 more close friends on average (hereinafter: all other demographic factors held constant).
Trust people in general (1 to 3)	0.121*	Community Life	Membership of a sport group is associated with the respondent being about 6% closer to reporting that “most people can be trusted” rather than that “you can’t be too careful dealing with people”. This can also be interpreted as a 12% higher likelihood of a mild trust increase (one step higher on the 3-point trust scale).
Feel belonging to neighbourhood (1 to 5)	0.075***	Understanding Society	Membership of a sport group is associated with the respondent being about 1.9% closer to strongly agreeing that he/she belongs to his/her neighbourhood (rather than strongly disagreeing) or a 7.5% likelihood of a one point positive increase in the feeling of belonging.
Volunteered in the last 12 months (0/1)	0.077***	Community Life	Sport group participation is associated with a 7.7 percentage points higher likelihood of volunteering.
Can achieve own goals (1 to 5)	0.122***	Active Lives	Membership of a sport group is associated with the respondent being about 3% closer to strongly agreeing that he/she can achieve the goals he/she sets him/herself (rather than strongly disagreeing). This can also be interpreted as a 12% higher likelihood of being one step higher on the agreement scale.
Life satisfaction (1 to 7 or 0 to 10)	0.074***	Understanding Society	Sport group participation is associated with being about 1.2% closer to being completely satisfied with life as opposed to absolutely miserable. Alternatively we can say that it has 7.4% higher likelihood of having a positive increase in life satisfaction (1 point on the scale). The association with life satisfaction is comparable to that of being religious.
General health (1 to 5)	0.270***	Understanding Society	Sport group participation is associated with being about 6.8% closer to having excellent (self-perceived) health as opposed to poor health, or a 27% chance of a smaller health increase.

Notes: OLS regressions. Coefficients of ‘sports club membership’ variable shown. Most outcomes are ordinal variables, that is, expressed on subjective scales such as 1 (strongly disagree) to 5 (strongly agree). The regression coefficient represents how much farther along the scale sport club members as opposed to non-members. For ease of interpretation, it has been expressed as a percentage of the total range of the scale that the variable uses. In this way, one can say that sport group participants stand x% closer to the positive end of the (e.g.) trust spectrum than non-participants. All models include control variables for a wide range of determinants of health and wellbeing as set out in Fujiwara and Campbell (2011). List of control variables and their coefficients provided in the Methodology section. Stars indicate statistical significance levels: *** < 1%; ** < 5%; * < 10% significance. Heteroskedasticity-robust standard errors used.

These findings are correlative (it may be that happier/healthier people select into sports clubs) and therefore the results, although statistically significant, do not indicate causality. Causality can only be inferred by designing Randomised Control Trials, which are outside of the scope of this study.

We apply best practice quasi-experimental methods for dealing with observational data of this type, by controlling for (holding constant) a range of potentially confounding demographic factors that may drive the outcomes like socio-economic status, religion, and geographical location. Holding these factors constant, the findings are consistent across all datasets which gives convergent validity to our results.⁶

The evidence is favourable and supports our hypothesis that sport group participation is beneficial for a young person’s wellbeing and social capital, drawing from a diverse pool of nationally representative UK household surveys.

The data also suggests that these benefits of sport group engagement are greater, across the board, for lower socio-economic groups.

A key aspect of this study is to investigate whether the potential benefits of membership in sporting groups are stronger or weaker for different subpopulations, including vulnerable groups in particular. For this purpose we ran a modified model which allows us to identify associations of different intensity for higher and lower socio-economic groups, as well as for rural and urban residents and by region.

The main insights are presented in the tables below. Where we see asterix (*) in the table, this suggests a positive and significant finding, with three stars (***) being the most positive and significant.

What we can observe is that most, but not all, of the wellbeing and social outcomes and benefits are magnified with sport club membership for lower socio-economic groups.

Tables 1B-1C. Association between sport groups and outcomes by socio-economic class (Active Lives dataset).

Active Lives	Can achieve goals	Trust people in local area	Volunteering	Life satisfaction	BMI
Higher SEG	0.077	0.079	0.264***	0.176*	-0.232
Lower SEG	0.192***	-0.023	0.435***	0.545***	-0.355

⁶ To shed more light on causality (that is, to identify whether membership in a sport club/organisation makes you more socially active or people who are initially more socially active self-select into sports clubs), an interesting area for future research would be to generate data that allows for more robust estimation techniques, such as fixed effects analysis. This would require conducting or identifying a longitudinal survey which asked questions about sport club membership and social outcomes of the same respondents for at least several years.

Community Life	Trust people in general	Trust neighbours	Volunteering	Formal volunteering	Informal volunteering
Higher SEG	0.073	0.019	0.055	0.039	0.044
Lower SEG	0.033	0.189**	0.090**	0.147***	0.012

Notes: OLS regressions. Coefficients of 'sports club membership' interacted with the variable in the row headers are shown. All models include control variables for a wide range of determinants of health and wellbeing as set out in Fujiwara and Campbell (2011) and described in Section 3.2.4. Stars indicate statistical significance levels: *** < 1%; ** < 5%; * < 10% significance. Heteroskedasticity-robust standard errors used.

Although there is some variation across datasets, the key findings identified above generally hold for most of the nationally representative UK datasets used in this study.

A detailed explanation of the methodology can be found in Section 2, with the full range of results of the work following in Section 3. Section 4 concludes.

2. Methodology

2.1. datasets used

We use five nationally representative UK datasets to investigate the relationship between wellbeing / social capital on one side and sport club membership on the other:

- The **Understanding Society** dataset is the successor of the discontinued British Household Panel Survey (BHPS). It collects a range of demographic and lifestyle data from individuals and families. The advantage of this dataset lies in its panel component – it is the only dataset where a significant number of respondents are sampled repeatedly in order to track their changes over time.
- The **Community Life** dataset is an annual cross-sectional survey collected by the UK Cabinet Office since 2012 to look at the latest trends in areas such as volunteering, charitable giving, local action and networks and wellbeing. Among the variables collected is participation in sport/exercise groups.
- The **Taking Part** dataset is a nationally representative database commissioned annually by the Department for Culture, Media and Sport along with partners at the Arts Council England, Historic England and Sport England. The survey collects data on aspects of leisure, culture and sport in England, in addition to the usual socio-demographic information on respondents.
- The **Active Lives** dataset is a Sport England-led survey about people in England (aged 14+) and their participation in leisure and recreational activities, including sport, physical activity and culture. The survey has an impressive sample, large enough to be representative of the adult population in each local authority area in England.
- The **Understanding Society Youth** dataset is collected alongside the main Understanding Society dataset using a separate questionnaire designed for children and teenagers aged 10-15. It collects a more limited, but still relevant set of demographics and social outcomes, and includes sport participation outside school.

It is important to note that the key variables of interest for the study – sport group participation and social outcomes – are often not the main topics of interest in some of these nationally representative datasets, and therefore are not collected in every wave or every questionnaire group. Below we list for each dataset the waves and sample size that contains the necessary information for this study, both for all age groups and for our target group of young people aged 25 or below:

Table 2. Datasets used for this study

Dataset name	Years (waves) Used	Total sample ⁷	Target age group	Target age sample	Of which sport group participants
Understanding Society (USoc)	2010/11, 2011/12, 2013/14, 2014/15 (waves 2, 3, 5, 6)	140,845	16-25	21,573	5,706
USoc Youth	2010/11, 2012/13, 2014/15 (waves 2, 4, 6)	11,929	10-15	11,929	6,095

⁷ Number of respondents for whom all the relevant variables (sport group participation, social outcomes, wellbeing and main demographics) are observed.

Taking Part	2012/13, 2013/14, 2014/15, 2015/16 (waves 8, 9, 10, 11)	27,447	16-25	2,932	454
Community Life	2012/13, 2013/14, 2014/15, 2015/16, 2016/17 (waves 1, 2, 3, 4, 5)	36,330	16-24	2,692	1,242
Active Lives	2016/17 (wave 2)	107,469	14-25 ⁸	12,073	4,931

2.2. Variables used in this analysis

In this section we will list the main variables used for our analysis, describing in more detail the form in which they appear in each of the datasets.

2.2.1. Sport club membership

Membership of a sport group, club or organisation (as defined within each dataset) is the main treatment variable of this study, that is, the main variable whose benefits we are trying to identify. The table below lists the respective questions (in each dataset) which we have used to flag whether the respondent participates in sport groups.

Table 3A. Sports club membership variables by dataset

Dataset name	Questions used to derive sport group participation	Response scale
Understanding Society (USoc)	<ul style="list-style-type: none"> - Are you currently a member of a sports club to do this sport / one of these sports? - waves 2 and 5 - Which organisations are you a member of? - "Sports club" is selected - waves 3 and 6 - Whether you are a member or not, do you join in the activities of any of these organisations on a regular basis? - "Sports club" is selected - waves 3 and 6 	All binary (yes/no). Waves 3/6 codes as "yes" if either of the two questions is a "yes."
USoc Youth	<ul style="list-style-type: none"> - Which of the following regular classes do you do outside school, if any? - "sport" is selected 	Binary (yes/no)
Taking Part	<ul style="list-style-type: none"> - Just thinking about the last 4 weeks, have you been a member of a club, particularly so that you can participate in any sports or recreational physical activities? - (if yes) What type of club was it? - "sports club" is selected 	Binary (yes/no), coded as "yes" if both questions are "yes."
Community Life	<ul style="list-style-type: none"> - Please pick out the ones which best describe any groups, clubs or organisations you've taken part in, supported or helped over the last 12 months. - "Sport/exercise (taking part, coaching or going to watch)" is selected 	Binary (yes/no)
Active Lives	<ul style="list-style-type: none"> - Are you currently a member of a club for the given activity? - A separate variable is coded for each sport and then aggregated to construct an indicator for club participation in ANY sport 	Binary (yes/no)

⁸ Respondents aged 14 and 15 are only included in some descriptive statistics, but excluded from regressions because they answered a different questionnaire.

2.2.2. Wellbeing outcomes

There are a range of variables which measure the respondent's subjective (self-perceived) wellbeing and general health. Among these are the four Office for National Statistics (ONS) wellbeing measures, which are collected in most of the datasets used in this study. A detailed list of the wellbeing measures by dataset can be found below:

Table 3B. Wellbeing and health variables by dataset

Wellbeing variable	Question form	Response scale	Datasets used in
Life satisfaction	How satisfied are you with your life nowadays?	1 to 7 (USoc and USoc Youth), 0 to 10 (all other datasets)	All datasets
Happiness	How happy did you feel yesterday?	0 to 10	Taking Part, Community Life, Active Lives
Anxiety	How anxious did you feel yesterday?	0 to 10	Taking Part, Community Life, Active Lives
Worthwhile life	To what extent do you feel the things you do in your life are worthwhile?	0 to 10	Taking Part, Community Life, Active Lives
General health	In general, would you say your health is... (multiple choice)	1 (poor) to 5 (excellent)	All except Active Lives
GHQ index	A sum of 12 mental health questions	Each question is 0 (no problems) to 3 (serious problems), resulting in a total of 0 to 36, with 0 being the best possible mental health state and 36 being the worst	Understanding Society
Body Mass Index (BMI)	Derived from the respondent's self-reported weight and height: weight divided by height squared.	Continuous (in kg/m ²), but categorical version also available, ranging from 1 (underweight) to 5 (morbidly obese)	Active Lives

2.2.3. Trust and community development / social capital outcomes

The aim of this study is to gather evidence in support of a benefit of sport group participation on social capital and community outcomes. The datasets mentioned above contain a range of variables that convey some information about this. These variables cover topics such as trust, volunteering, number and quality of friendships, affinity to one's community, closeness with one's family. Wherever appropriate, the scales have been inverted so that a higher value indicates a more desirable/positive outcome. A list of the key social outcomes present in different datasets and analysed in this study can be seen below:

Table 3C. Example selection of Community Development outcomes by dataset

Social outcome	Question form	Response scale (after recoding)	Datasets used in
Has friends	Do you have any friends?	Binary (yes/no)	USoc
Number of close friends	How many close friends would you say you have?	1, 2, 3, 4, 5, 6 (6-9), 7 (10-19), 8 (20+)	USoc, USoc Youth

Can rely on friends	How much can you rely on [your friends] if you have a serious problem?	1 (not at all) to 4 (a lot)	USoc
Loneliness	How often do you feel lonely? (Note that the scale has not been inverted to avoid confusing interpretations)	1 (never) to 5 (often/always)	Community Life
Trust in people in neighbourhood	Do you believe most people in your neighbourhood / local area can be trusted? (varies slightly by dataset)	1 (strongly disagree) to 5 (strongly agree)	USoc, Community Life, Active Lives
Trust in people in general	Would you say that most people can be trusted (3) or that you can't be too careful in dealing with people (1)? (2 - "it depends")	1 to 3	Community Life
Trust in institutions	How much do you trust the following: Parliament Your local council The police	1 (not at all) to 4 (a lot) for each	Community Life
Talking to people in neighbourhood	I regularly stop and talk with people in my neighbourhood.	1 (strongly disagree) to 5 (strongly agree)	USoc, Community Life
Belonging to neighbourhood	I feel like I belong to this neighbourhood (USoc) / How strongly do you feel you belong to your immediate neighbourhood (CL) / in your local area (TP)?	1 to 5 (USoc), 1 to 4 (TP, CL)	USoc, Taking Part, Community Life
Ethnic mixing in local area	To what extent do you agree or disagree that this local area is a place where people from different backgrounds get on well together?	1 (definitely disagree) to 4 (definitely agree)	Taking Part, Community Life
Volunteering in the last 12 months	In the last 12 months, have you given any unpaid help or worked as a volunteer for any type of local, national or international organisation or charity? (with slight variations by dataset)	Binary (yes/no)	All datasets
Formal volunteering in the last 12 months	Variable is derived from the answer to the main volunteering question and the name/type of the organisation with which the respondent volunteered	Binary (yes/no)	Community Life
Informal volunteering in the last 12 months	Variable is derived from the answer to the main volunteering question and the name/type of the organisation with which the respondent volunteered	Binary (yes/no)	Community Life

2.2.4. Demographic control variables

The purpose of including these variables in the analysis is to control for the main determinants of our wellbeing and social outcomes other than sport participation. Fujiwara and Campbell (2011)⁹ draw up a list of the main determinants of life satisfaction found in the literature to date, of which we try to include as much as each dataset provides. It is reasonable to believe that these factors are also likely to influence social outcomes such as trust or friendships. Furthermore, demographics are of interest in themselves in order to describe the composition of sport group participants and non-participants comparatively, and

⁹ Fujiwara, D. and Campbell, R. (2011). *Valuation Techniques for Social Cost-Benefit Analysis: Stated Preference, Revealed Preference and Subjective Well-Being Approaches. A Discussion of the Current Issues*. HM Treasury and DWP Joint Publication.

to paint a demographic profile of “the average club-goer.” A list of demographic variables used in this study and their availability by dataset can be found underneath:

Table 3D. Demographic control variables by dataset

Variable	Response scale / categories	Notes / Comments	Availability by dataset				
			USoc	USocY	TP	CL	AL
Age	Whole numbers indicating age in years	CL stopped collecting exact age in the online survey and only collects age bins. Our entire target group falls into the 16-24 age bin and hence we cannot control for age.	V	V	V		V
Gender	Male, Female		V	V	V	V	V
Income	Continuous (natural logarithm)	USoc has exact income. TP and CL have income bins, where we impute income as the midpoint. We then take the natural logarithm in all cases. TP has personal income; the rest have household income. AL has no income data.	V	V	V	V	
Marital status	Single, Married or civil partner, Separated, Divorced, Widowed	AL has a somewhat different variable indicating the type of household (single / couple / lone parent etc.)	V		V	V	V
Number of children in household	0, 1, 2, 3, 4+	4 or more children grouped together to avoid small bin sizes. Usually defined as children aged 0-15 in household, but USocY uses 0-13 and some datasets don't specify the age.	V	V	V	V	V
Number of adults in household	0, 1, 2, 3, 4, 5, 6, 7+	OECD definition used: people aged 14+ are considered adults. Used for youth dataset only.		V			
Education	Degree or above, Other higher education, A-levels, GCSE, No qualifications	Varies by dataset. CL includes postgraduate education as a separate category. AL bundles all level 4 education (Certificate of HE and above) together.	V		V	V	V
Employment status	Employed, Unemployed, Student, Retired etc.	Varies by dataset. Some distinguish between full-time, part-time employment and self employment. Others group everyone not in the labour force together.	V		V	V	V
Rural / urban area	Rural, Urban		V	V	V	V	V
Respondent is religious	Religious, not religious	The wellbeing literature doesn't recommend controlling for individual faiths. We cannot control for religion in AL or USocY	V		V	V	

		because it is not collected in the same wave or questionnaire group as the outcomes.					
Ethnicity (broad categories)	White, Mixed, Asian, Black, Other	In USocY not collected in the same wave as sport group participation.	V		V	V	V
Disability	No disability, Non-limiting disab., Limiting disab.	Active Lives only.					V
Socioeconomic class (job-based)	Higher SEC, Lower SEC, SEC residuals	In most datasets, this is based on the National Statistics Socio-Economic Classification (NS-SEC), which we condense into three categories (Higher for NS-SEC 1-4, Lower for NS-SEC 5-8, residuals for NS-SEC 9). However, USoc uses a different nomenclature.	V		V	V	V
Socioeconomic class (income-based)	Above median income, Below median income	An alternative measure of SEC derived by comparing the respondent's income (household or personal, depending on what is available) to the median income in the dataset.	V	V	V	V	
Immigration status	Native, First-generation immigrant, Second-gen. imm.	Only used for heterogeneous effects analysis. Not used as a control variable	V				
LSOA Index of Multiple Deprivation	1 (most deprived) to 10 (least deprived)	A ranking of all LSOAs in England by their level of deprivation, grouped into deciles.			V	V	V
Region	9 regions of England	USoc also has Scotland, Wales and Northern Ireland, each as one separate extra category.	V	V	V	V	V
Carer status	Yes, No	Indicates whether the respondent has to take care of someone in the household.	V				
House ownership	Owned outright, Mortgage, Rented, Rent-free, Other		V			V	
Wave of survey	dataset dependent	Indicator for each wave in the dataset. Included to account for time trends.	V	V	V	V	
Survey mode	Face-to-face, Online, Postal	CL only. However, postal surveys are excluded from the regression because they do not collect all relevant variables. Active Lives is online only (there is a small postal component but it does not collect					V

		all required data). All the other surveys are face-to-face only.				
Interview month	January to December	Included to account for seasonality.	V	V	V	V

General health (presented in the wellbeing outcomes table earlier) is also used as a control variable when other outcomes are analysed as it is also an important determinant of wellbeing and social connections.

Note on socio-economic class:

Socio-economic class is a key element of this study, given that the main narrative points to a trust and social capital deficit among lower socio-economic classes. There are two measures of socio-economic class used in this study. The first is based on the respondent’s job / occupation. The second is based on how the respondent’s income compares to that of other respondents in the data, where most young people fall into the lower income category (because one’s earnings at this stage of life are usually lower). Also in all relevant analysis we control for wider ‘household income’ in the data so we can be confident we are capturing socio-economic characteristics in the analysis.

In the job-based socio-economic stratification, we have THREE categories that are considered in the regression:

- Higher SEG, which groups NS-SEC categories 1-4 (managerial, professional, administrative, intermediate occupations, small employers and own account workers)
- Lower SEG, which groups NS-SEC categories 5-8 (lower supervisory and technical, semi-routine, routine occupations and the long-term unemployed)
- Those who are not classified because they are not in the labour force. This includes students, retired, disabled, stay-at-home people and others.

Most respondents in the target group of young people aged under 25 fall into the third category usually because they are students. In the disaggregated analysis, a distinct association between each outcome and sport group participation is estimated for each of the above three categories via regression models with interaction terms. This ensures that everyone’s occupation class and labour force status is accounted for, and it also implies that the higher vs. lower SEG differential we talk about throughout this paper refers to those young people who have a job or are looking for a job.

Note on religion, household income and trust in the Active Lives dataset:

We can see in the table 3D above that the Active Lives dataset does not enable controls for household income and religion. Religious upbringing and regular attendance are significantly correlated with higher trust scores, so controlling for religion when reporting trust outcomes is important¹⁰. This requires us to treat the most positive outcomes from Active Lives in the tables further below with some caution when comparing them to the other datasets that do

¹⁰ <https://ourworldindata.org/trust> Taking these results at face value, the reported effects suggest that being raised religiously raises the level of trust by 2.6 percent; and regularly attending religious services (the author's definition of being "religious" for the purpose of the figure), raises the level of trust by another 20 percent.

contain these controls. In short, the Active Lives results may be more positive than would be the case if we were able to control for the income and religion variables.

2.3. The model

The analysis starts off by tabulating wellbeing and social outcomes as well as demographics for sport group participants and non-participants. However, while this may point out that the outcomes and sport group participation move in the same direction, it is insufficient to ascertain whether this indicates a benefit that membership in sports clubs has for one's wellbeing or social capital, or rather is an artefact of selection. It could be the case that more well-off or more socially-connected respondents are more likely to engage in sport groups for other reasons, be it because they are richer, younger or more educated (for instance).

This leads us to move on to regression analysis, where we can control for the many determinants of wellbeing / social outcomes that we listed in the "Control variables" section above. This is achieved by including them in an Ordinary Least Squares (OLS) regression equation such as the following:

$$O_i = \alpha + \beta_1 Sp_i + \beta_2 X_i + \varepsilon_i \quad (1)$$

Here, O_i is the outcome variable, as listed in the "Wellbeing outcomes" and "Social capital outcomes" subsections; a separate regression is run for each outcome variable in each dataset. Sp_i is a dummy variable indicating whether the respondent participates in a sport club or organisation, as described in the "Sport club membership" subsection for the respective dataset. X_i is a vector of all the control variables available in the dataset, as listed in the "Demographic controls" subsection; and ε is the error term containing *unobserved* factors that determine the relevant outcomes.¹¹

To further insure against the influence of *unobserved* factors (a.k.a. omitted variable bias), we supplement OLS analysis by Fixed Effects regressions in Understanding Society, which is a panel dataset. The FE model is similar to OLS but looks only at the changes that occur between waves in the variables of an individual that was surveyed more than once. Fixed effects generally confirm OLS results but mostly lack statistical significance. This is due to the poor panel structure even of the USoc data – given that our key variables are usually only collected every other wave, this leaves an average number of observations per individual of only 1.5 – not enough to conduct robust panel data analysis. We therefore fall back to OLS results as our main findings and supplement these with the caveat that the regression coefficient only represents an association between sport group participation and the wellbeing / social outcome, holding a range of demographic factors constant, but does not

¹¹ There are a series of technical assumptions which underpin the validity of OLS regressions. Among these are random sampling (which the nationally representative surveys do their best to ensure), a true linear relationship between the variables (which can be circumvented to allow for a more flexible functional form by using an age squared term and dummy variables for categorical controls), and the absence of other factors which influence outcomes and sport group participation at the same time. The last assumption is the most difficult to verify, as there are a plethora of factors which can determine wellbeing and sport.

necessarily indicate a causal effect or a direct benefit of sport groups on wellbeing / social capital.

Furthermore, we can investigate how the association between outcomes and sport group participation varies by socio-economic class, region, urban/rural residence, and immigration status (the latter is available in USoc only). This is done with the help of regression models with interaction terms, such as Equation 2 below (where SEG_i is a socio-economic group dummy). This would feed into the objective of identifying whether sport group participation is associated with higher improvements in social outcomes or wellbeing for disadvantaged groups.

$$O_i = \alpha + \beta_1 Sp_i + \beta_2 Sp_i SEG_i + \beta_3 X_i + \varepsilon_i \quad (2)$$

The categorical outcome variables, which are coded on ordinal (0-to-10, 1-to-7, 1-to-5 or other) scales, will be treated as cardinal for the purpose of this analysis. Kristoffersen (2015) shows that the cardinality assumption is reasonable in most research contexts¹², and at the same time it facilitates interpretation and even subsequent monetary valuation of the results.

¹²

https://melbourneinstitute.unimelb.edu.au/assets/documents/hilda-bibliography/working-discussion-research-papers/2011/Kristoffersen,-I_-The-Subjective-Wellbeing-Scale-How-Reasonable-is-the-Cardinality-Assumption_dp11.15.pdf

3. Results in full

3.1. The trust and social cohesion deficit in UK society

There is extensive evidence pointing to the fact that lower socio-economic groups in the UK are considerably less trusting of their neighbours and of people in general. For example, in one of our previous reports, “A Bit Rich: Why is volunteering biased towards higher socio-economic groups?” (published in January 2019), analysis of the Community Life data shows that lower SEG in the national sample have significantly lower odds of having mixed with people from different ethnic groups at least once in the last 12 months (25% of the sample, compared to 37% of those from higher SEG) and also report slightly lower levels of trust in people in the local neighbourhood (40% agree that they trust many people, compared to the national average of 45%, and 51% with higher SEG). This is detailed in Table 4A below. There are also some interesting divergences in the trust that different groups hold for government institutions at the local and national level.

The Community Life dataset provides a number of variables related to trust in social institutions (local council, parliament, and the police), although these questions are only asked in one wave of Community Life (2012/13), which reduces the sample size, and therefore the strength of the statistical analysis. We also measure overall trust through a question on trust in the neighbourhood. In combination, these variables provide an important indication of whether the levels of trust within lower SEG are driven by (lack of) trust in official institutions, or (lack of) trust in the community.

Table 4A. Trust among lower socio-economic groups: Community Life

	Lower SEG Odds ratio	Lower SEG (%, n/N)	Higher SEG (%, n/N)	Total observations (%, n/N)
Trust local council: a lot or a fair amount	1.070	63.5% (2221/3499)	61.1% (1590/2602)	62.5% (4181/6692)
Trust Parliament: a lot or a fair amount	0.953	31.4% (1107/3526)	32.9% (871/2647)	31.7% (2147/6773)
Trust police: a lot or a fair amount	0.867*	82.6% (2975/3602)	86.7% (2305/2658)	84.1% (5789/6882)
Trust many of the people in neighbourhood	0.771***	40.1% (3074/7668)	50.5% (3700/7322)	44.7% (7496/16754)

*Notes: Community Life 2012-2016, England. Lower socioeconomic group defined as below sample median household income. Column 2: Logistic regression model. Heteroskedasticity-robust standard errors used. *** <1% significance; ** <5% significance; * <10% significance. Standard socio-demographic controls including household income (ethnic group regressions only), employment status, marital status, geographical region etc. included but not reported in this table. Weights not applied due to sample size issues.*

To further our evidence on the ‘trust gap’ between higher and lower socio-economic groups, **in this study** we investigated a number of nationally representative UK datasets with the focus on young people aged below 25, which is our target group. We can see in the tables below that young people from lower SEG indeed tend to be consistently less trusting of their neighbours, although they talk to them more. They are also less socially connected and their pool of friends is less diverse (at least according to Understanding Society data, with a less

clear pattern for Community Life). Furthermore, lower SEG are considerably less likely to volunteer compared to higher SEG.

Table 4B. Trust and social mixing by socio-economic group: Understanding Society, age 16-25

	Higher SEG ¹³	Lower SEG	SEG unclassified	Total Observations
Number of observations	5426	20087	25147	50660
I trust people in this neighbourhood (1 to 5)	3.44	3.4	3.35	3.38
I talk regularly to neighbors (1 to 5)	3.07	3.27	3.21	3.22
Number of close friends of respondent, binned (1 to 8)	4.65	4.32	4.14	4.27
Volunteered in the last 12 months	22.1% (434/1961)	19.2% (1440/7492)	22.7% (2025/8935)	21.2% (3899/18388)
Friends have diverse ages (1 to 4)	1.96	1.88	1.75	1.83
Friends have diverse race (1 to 4)	1.93	1.84	2.01	1.93
Friends have diverse education (1 to 4)	2.09	2	1.94	1.98
Friends live in different areas (1 to 4)	3.2	2.96	2.97	2.99

Table 4C. Trust and social outcomes by socio-economic group: Active Lives, age 14-25

	Higher SEG	Lower SEG	SEG unclassified	Total Observations
Number of observations	3146	1905	7022	12073
I can achieve most goals I set myself (1 to 5)	3.9	3.63	3.84	3.82
Most of the people in my local area can be trusted (1 to 5)	3.13	3.01	3.16	3.12
Volunteered in the last 12 months to support sport and physical activity	23.5% (361/1538)	19.1% (186/975)	34.3% (1662/4846)	30.0% (2209/7359)

Table 4D. Trust and social outcomes by socio-economic group: Community Life, age 16-24

	Higher SEG	Lower SEG	SEG unclassified	Total Observations
Number of observations	628	798	1750	3176
Trusts people in neighbourhood (1 to 4)	2.92	2.73	2.89	2.86
Chats to people in neighbourhood (1 to 5)	2.64	3	2.72	2.77
Trusts people in general (1 to 3)	1.92	1.7	1.85	1.83
Trust in: The police (1 to 4)	3.12	2.91	3.07	3.03
Trust in: Parliament (1 to 4)	2.4	2.02	2.41	2.27
Trust in: The local council (1 to 4)	2.8	2.66	2.78	2.74
Mixing with friends of different race (1 to 4)	1.97	1.97	2.23	2.1
Mixing with friends of different faith (1 to 4)	2.14	2.06	2.39	2.25
Mixing with friends of different age (1 to 4)	2.17	2.16	1.86	2.01
Mixing with friends of diff. education (1 to 4)	2.2	2.29	1.9	2.06

¹³ Lower socio-economic group defined as working in lower supervisory and technical occupations, semi-routine occupations, routine occupations, as well as those who never worked and the long-term unemployed (NS-SEG 5-8). Higher SEG: NS-SEC 1-4. SEG unclassified: students, retired and others not in the labour force (NS-SEG 9).

3.2. The role of sport in addressing trust and social cohesion in society

While in the previous section we have seen that lower socio-economic groups have lower levels of trust and social capital to begin with, in this section we will see how participation in a sport club or organisation can offer a potential solution to address this trust gap / deficit. This will be done in three steps, each supported by evidence from the numerous nationally representative datasets analysed as part of this study.

First of all, the descriptive statistics (Tables 5A-5E) present how sports club/organisation members generally have higher levels of wellbeing, trust and social capital than non-members. This finding is taken further by the regression analysis results (Table 6), which show that the positive association between sport club membership and these outcomes persists even after controlling for demographic factors which are likely to be responsible for the selection of happier or more social people into sport clubs. Finally, disaggregated analysis by socio-economic group shows that the positive association between sport club membership and most of these outcomes is stronger for lower SEG respondents (Tables 7A-7D). Thus we show that participation in sport organisations is a positive option to bridge the deficit in trust and social capital between the more affluent and more vulnerable groups in society.

3.2.1. Descriptive statistics

The tables below compare the main wellbeing and social outcomes, as well as the demographic composition of sports club members and non-members in the Understanding Society, Taking Part, Community Life, Active Lives and Understanding Society Youth datasets. The comparison is made for the target subsample of young respondents aged 25 or below.

In the Understanding Society data we can see that sports club members are happier and healthier, both in terms of self-perceived general physical health as well as mental health. They have only slightly higher proportions of friends of different ethnicities, but not age or education. However, their overall number of friends is considerably higher and they are less likely to have no friends at all. They have higher levels of trust in their neighbours and overall a stronger bond to their local communities. They are more likely to volunteer. Demographically they tend to be younger, richer, more predominantly male, without children, more educated and white.

Table 5A. Wellbeing, social outcomes and demographics by sports club membership in the Understanding Society panel dataset, waves 2, 3, 5, 6

Variable	Age 16 to 25		
	Not a sports club member	Sports club member	Total
Sample size	15867	5706	21573
Wellbeing outcomes			
Life satisfaction (1 to 7)	5.21	5.48	5.28
General health (1 to 5)	3.71	4.07	3.81

Mental health problems - GHQ index, 0(best) to 36(worst)	11.06	10	10.78
Social, Mixing and Community Outcomes			
Friends have diverse ages (1 to 4), mean	1.85	1.77	1.83
Friends have diverse race (1 to 4), mean	1.92	1.94	1.92
Friends have diverse education (1 to 4), mean	2	1.92	1.98
Friends live in different areas (1 to 4), mean	3.01	2.93	2.99
Has no friends	2.0% (139/6905)	0.7% (18/2599)	1.7% (157/9504)
Number of close friends of respondent, binned (1 to 8), mean	4.11	4.71	4.26
I can rely upon my friends (1 to 4), mean	3.3	3.38	3.32
I trust people in this neighbourhood. (1 to 5), mean	3.34	3.5	3.38
I talk regularly to neighbours (1 to 5), mean	3.18	3.27	3.21
I feel I belong to this neighbourhood. (1 to 5), mean	3.48	3.6	3.51
The friendships in my neighbourhood mean a lot to me (1 to 5), mean	3.28	3.39	3.31
Volunteered in the last 12 months	21.5% (1666/7745)	27.6% (786/2847)	23.1% (2452/10592)
Hours volunteered in the last 4 weeks, mean	1.98	2.21	2.04
Demographics			
Household income ¹⁴ (monthly)	£3,845	£4,611	£4,048
Age	20.2	19.92	20.12
Female	58.2% (9227/15867)	36.9% (2106/5706)	52.5% (11333/21573)
Married	3.9% (615/15831)	1.9% (109/5688)	3.4% (724/21519)
No children	89.0% (14126/15867)	96.3% (5495/5706)	91.0% (19621/21573)
Higher education degree	12.8% (1985/15471)	14.7% (816/5555)	13.3% (2801/21026)
Employed (full or part-time)	38.2% (6034/15792)	38.8% (2199/5673)	38.4% (8233/21465)
Urban area	80.8% (12795/15839)	77.5% (4417/5697)	79.9% (17212/21536)
Religious	43.9% (4578/10422)	43.0% (1503/3494)	43.7% (6081/13916)
White	72.6% (11523/15867)	76.2% (4347/5706)	73.6% (15870/21573)
Current job socio-economic classification is low ¹⁵	41.1% (6485/15772)	42.2% (2396/5679)	41.4% (8881/21451)
Equivalised household income is below median	58.5% (9244/15815)	44.9% (2551/5684)	54.9% (11795/21499)
Immigration status: native	97.0% (15392/15867)	97.3% (5552/5706)	97.1% (20944/21573)

Notes: The statistics calculated above exclude those respondents for whom the variable of interest (in the row header) is unknown.

¹⁴ Contains the earnings of everyone living in that household (including parents if respondent lives with his/her parents) from all sources of income.

¹⁵ Note that full-time students are outside the NS-SEC classification, making up the so-called NS-SEC residuals. Because the share of students is high among the young subsample, the share of both high and low socio-economic status according to NS-SEC is negatively affected.

The Understanding Society Youth survey further investigates some social outcomes in children and adolescents aged 10 to 15. The question coming closest to identifying group sport participation is “does regular sports classes outside school.” Children who do regular sports classes outside school are happier, healthier, have more friends (and more satisfying friendships), spend more time with their families and volunteer more. They also score higher on all self-confidence outcomes (and lower on the measures indicating lack of confidence, respectively).

Again, they tend to come from richer families, are younger and more predominantly male, and are more likely to come from rural areas. They have slightly more siblings but fewer adults in the household.

Table 5B. Wellbeing, social outcomes and demographics by sports activities after class in the Understanding Society Youth panel dataset, waves 2, 4, 6

Does regular sports classes outside school	Age 10 to 15		
	No	Yes	Total
Sample Size	5834	6095	11929
Wellbeing outcomes			
Life satisfaction (1 to 7)	5.77	6.06	5.92
General health (1 to 5)	3.64	3.98	3.81
Social, Mixing and Community Outcomes			
Number of close friends, binned (1 to 8), mean	4.42	4.85	4.64
Satisfaction with friends (1 to 7), mean	6.23	6.37	6.3
Frequency of eating evening meals with family (1 to 4), mean	3.09	3.21	3.15
Volunteers more than once a year (0/1), mean	0.33	0.40	0.37
Frequency of volunteering (1 to 6), mean	2.16	2.40	2.28
Self-confidence outcomes (all 1 to 4), means			
I feel I have a number of good qualities	3.17	3.32	3.25
I don't have much to be proud of	2.01	1.81	1.91
I certainly feel useless at times	2.32	2.17	2.24
I am as able as most people	3.14	3.27	3.21
I am a likeable person	3.24	3.34	3.29
I can usually solve my own problems	3.08	3.15	3.12
I am inclined to feel I am a failure	1.68	1.54	1.61
At times I feel I am no good at all	2.06	1.88	1.97
Demographics			
Household income (monthly)	£3,690	£4,350	£4,027
Age	12.71	12.37	12.54
Number of adults aged 14+ in household (OECD definition)	2.85	2.77	2.81

Number of children aged 0-13 in household, mean	1.58	1.66	1.62
Female	62.8% (3661/5834)	36.4% (2221/6095)	49.3% (5882/11929)
Urban area	79.6% (4637/5826)	73.3% (4461/6086)	76.4% (9098/11912)
Equivalised household income is below median	55.5% (3229/5817)	43.0% (2611/6075)	49.1% (5840/11892)

Notes: The statistics calculated above exclude those respondents for whom the variable of interest (in the row header) is unknown.

The Active Lives dataset introduces the three other ONS wellbeing measures – happiness and a sense of worthwhile life, which are considerably higher for sports club members along with life satisfaction, and the inverted wellbeing measure of anxiety, which is accordingly lower. There is no general health indicator, but it can be proxied by the Body Mass Index, which is lower for sports club members, with a substantially lower share of obese respondents. Furthermore, sports club members score higher on both the individual and community ‘development’ indicators – perseverance and trust – and also (quite obviously) volunteer in sport much more frequently. Confirming the trend, they are younger, more likely to be male, single, without children, come from more affluent areas, and have slightly higher shares of rural and white respondents.

Table 5C. Wellbeing, social outcomes and demographics by sports club membership in the Active Lives dataset, wave 2

Variable	Age 14 to 25		
	Not a sports club member	Sports club member	Total
Sample size	7142	4931	12073
Wellbeing outcomes			
Life satisfaction, 0-to-10 scale, mean	6.65	7.22	6.85
Happiness, 0-to-10 scale, mean	6.71	7.15	6.86
Anxiety, 0-to-10 scale, mean	3.99	3.84	3.93
Worthwhile life, 0-to-10 scale, mean	6.74	7.24	6.91
Respondent's BMI, mean	23.82	22.87	23.46
Respondent's BMI, grouped (1 - lightest to 5 - heaviest) mean	2.34	2.18	2.28
Social, Mixing and Community Outcomes			
I can achieve most of the goals I set myself, mean	3.75	3.95	3.82
Most of the people in my local area can be trusted, mean	3.05	3.26	3.12
Volunteered in the last 12 months to support sport and physical activity, mean	0.15	0.49	0.3
Volunteered in sport more than once in the last 12 months, excluding raising fun, mean	0.08	0.38	0.22
Volunteered in sport in the last 4 weeks, excluding raising funds, mean	0.05	0.22	0.12
Demographics			

Age	19.95	18.25	19.25
Female	64.4% (4599/7142)	52.8% (2602/4928)	59.7% (7201/12070)
Lives as a couple	22.3% (1299/5824)	13.9% (450/3247)	19.3% (1749/9071)
No children	74.3% (5276/7105)	82.3% (4050/4919)	77.6% (9326/12024)
Level 4 education (Certificate of HE) or above	35.3% (2137/6048)	36.0% (1192/3313)	35.6% (3329/9361)
Employed (full or part-time)	44.3% (2685/6056)	37.0% (1225/3315)	41.7% (3910/9371)
Urban area	87.6% (6254/7142)	80.3% (3959/4931)	84.6% (10213/12073)
Religious	51.7% (1975/3821)	51.6% (1567/3038)	51.6% (3542/6859)
White	78.0% (5322/6825)	84.4% (4058/4810)	80.6% (9380/11635)
LSOA Index of Multiple Deprivation decile, mean	4.38	5.54	4.86
current job socio-economic classification is low ¹⁶	20.5% (1461/7142)	9.0% (444/4931)	15.8% (1905/12073)

Notes: The statistics calculated above exclude those respondents for whom the variable of interest (in the row header) is unknown.

The Taking Part dataset confirms the wellbeing differentials revealed in the Active Lives data and the health differential found in Understanding Society, as well as the demographic characteristics mentioned earlier. Few social outcomes can be correlated to sports club membership due to data collection issues – variables such as trust were not collected in the same wave as sport club participation. Those that were collected in the same waves are all higher for sports club members.

Table 5D. Wellbeing, social outcomes and demographics by sports club membership in the Taking Part dataset, waves 8-11

Variable	Age 16 to 25		
	Not a sports club member	Sports club member	Total
Sample size	2478	454	2932
Wellbeing outcomes			
Life satisfaction, 0-to-10 scale, mean	7.85	8.23	7.91
Happiness, 0-to-10 scale, mean	8	8.35	8.05
Anxiety, 0-to-10 scale, mean	2.67	2.53	2.65
Worthwhile life, 0-to-10 scale, mean	7.89	8.11	7.92
General health, 1-to-5, self-reported, mean	4.32	4.58	4.36
Social, Mixing and Community Outcomes			
Feels that he/she belongs to local area (1 to 4)	2.8	2.95	2.82
People from different backgrounds get on well together (1 to 4)	2.93	3.09	2.96
Spends time with family and friends	91.0% (2256/2478)	93.6% (425/454)	91.4% (2681/2932)
Volunteered in the last 12 months	28.7% (710/2475)	41.4% (187/452)	30.6% (897/2927)

¹⁶ Note that full-time students are outside the NS-SEC classification, making up the so-called NS-SEC residuals. Because the share of students is high among the young subsample, the share of both high and low socio-economic status according to NS-SEC is negatively affected.

Demographics			
Personal earnings (yearly)	£8,947	£9,425	£9,022
Age	20.56	19.48	20.4
Female	58.9% (1459/2478)	25.6% (116/454)	53.7% (1575/2932)
Married	4.2% (105/2477)	1.1% (5/454)	3.8% (110/2931)
No children	62.4% (1546/2478)	74.0% (336/454)	64.2% (1882/2932)
Higher education degree (including professional)	16.2% (401/2478)	15.6% (71/454)	16.1% (472/2932)
Doing paid work (includes self-employed)	54.2% (1344/2478)	57.3% (260/454)	54.7% (1604/2932)
Urban area	84.2% (2087/2478)	80.2% (364/454)	83.6% (2451/2932)
Religious	42.9% (1061/2474)	46.0% (208/452)	43.4% (1269/2926)
White	85.4% (2116/2477)	89.2% (405/454)	86.0% (2521/2931)
LSOA Index of Multiple Deprivation decile, mean	4.92	6.18	5.11
Low socio-economic class (NS-SEC 5-8) ¹⁷	31.9% (790/2478)	20.9% (95/454)	30.2% (885/2932)
Personal income is below median	65.4% (1539/2355)	64.4% (280/435)	65.2% (1819/2790)

Notes: The statistics calculated above exclude those respondents for whom the variable of interest (in the row header) is unknown.

The Community Life dataset is interesting in that it provides a series of extra social outcomes that are not present in the other datasets. For example, we can see that the higher likelihood to volunteer or sports club member is especially concentrated around formal volunteering. Furthermore, alongside higher levels of general trust in people and trust in neighbours, we can see that members of sport organisations also have higher levels of trust in institutions (the police, Parliament, the local council). They also report less loneliness and more satisfaction with the local area, but not necessarily more ethnic/religious mixing. Wellbeing differences and socio-demographic differences from previous datasets are confirmed.

Table 5E. Wellbeing, social outcomes and demographics by sports club membership in the Community Life dataset, waves 1-5

Variable	Age 16 to 24		
	Not a sports club member	Sports club member	Total
Sample size	1450	1242	2692
Wellbeing outcomes			
Life satisfaction, 0-to-10 scale, mean	6.97	7.34	7.15
Happiness, 0-to-10 scale, mean	6.78	7.29	7.02
Anxiety, 0-to-10 scale, mean	3.79	3.36	3.59
Worthwhile life, 0-to-10 scale, mean	7.04	7.41	7.21
General health, 1-to-5, self-reported, mean	4.1	4.3	4.19

¹⁷ Note that full-time students are outside the NS-SEC classification, making up the so-called NS-SEC residuals. Because the share of students is high among the young subsample, the share of both high and low socio-economic status according to NS-SEC is negatively affected.

Social, Mixing and Community Outcomes			
Formal or informal volunteering in the last 12 months	65.8% (954/1450)	80.5% (1000/1242)	72.6% (1954/2692)
Formal volunteering in last 12 months	38.2% (554/1450)	62.5% (776/1242)	49.4% (1330/2692)
Informal help in last 12 months	54.0% (783/1450)	61.1% (759/1242)	57.3% (1542/2692)
Ethnic or religious mixing during formal volunteering	27.3% (365/1336)	46.9% (567/1208)	36.6% (932/2544)
Trusts people in neighbourhood (1 to 4), mean	2.75	3.02	2.87
Chats to people in neighbourhood (1 to 5), mean	2.68	2.82	2.74
Trusts people in general (1 to 3), mean	1.8	1.95	1.87
Trust in: The police (1 to 4), mean	3.01	3.08	3.05
Trust in : Parliament (1 to 4), mean	2.28	2.37	2.33
Trust in: The local council (1 to 4), mean	2.73	2.81	2.78
How often meets family and friends (1 to 4), mean	2.94	3.08	3.01
How often respondent feels lonely (1 to 4), mean	3.17	3.02	3.11
Mixing with friends of different race (1 to 4), mean	2.12	2.14	2.13
Mixing with friends of different faith (1 to 4), mean	2.32	2.29	2.3
Mixing with friends of different age (1 to 4), mean	2.06	1.98	2.02
Mixing with friends of diff. education (1 to 4), mean	2.1	1.99	2.05
Can influence local area (1 to 4), mean	2.01	2.14	2.07
Respondent belongs in local area (1 to 4), mean	3.08	2.93	2.99
People from different backgrounds get on well together in local area (1 to 4), mean	2.91	3.02	2.96
Satisfaction with local area (1 to 5), mean	3.74	3.99	3.86
Demographics			
Face-to-face survey	23.1% (335/1450)	38.7% (481/1242)	30.3% (816/2692)
Household income (yearly)	£9,940	£10,747	£10,331
Age	20.28	20.1	20.19
Female	66.5% (963/1449)	49.6% (616/1242)	58.7% (1579/2691)
Married	4.1% (56/1382)	2.5% (30/1222)	3.3% (86/2604)
No children	65.3% (945/1448)	70.4% (874/1241)	67.6% (1819/2689)
Higher education degree	20.0% (260/1299)	22.7% (266/1172)	21.3% (526/2471)
Employed (full- or part-time or self-employed)	50.9% (661/1299)	56.6% (679/1200)	53.6% (1340/2499)
Urban area	89.5% (1160/1296)	85.6% (975/1139)	87.7% (2135/2435)
Religious	55.3% (798/1443)	55.1% (678/1231)	55.2% (1476/2674)
White	74.5% (1071/1438)	79.6% (982/1233)	76.9% (2053/2671)
Low socio-economic class (NS-SEC 5-8) ¹⁸	23.4% (339/1450)	18.1% (225/1242)	21.0% (564/2692)
Income is below median	83.6% (971/1162)	82.7% (903/1092)	83.1% (1874/2254)

¹⁸ Note that full-time students are outside the NS-SEC classification, making up the so-called NS-SEC residuals. Because the share of students is high among the young subsample, the share of both high and low socio-economic status according to NS-SEC is negatively affected.

LSOA Index of Multiple Deprivation, decile (1 - most deprived)	4.38	5.3	4.81
--	------	-----	------

Notes: The statistics calculated above exclude those respondents for whom the variable of interest (in the row header) is unknown.

All in all, we can see that sports club members report considerably **higher levels of wellbeing and health** than non-members. This is true for all available wellbeing measures – life satisfaction and general health, mental health (USoc), BMI/obesity (Active Lives), happiness, anxiety and a sense of worthwhile life (CL, TP, Active Lives).

In terms of social outcomes, a pervasive finding throughout all the datasets is that sports club members are **more socially connected** – they report having more close friends, relying on their friends more, and spending more time with their family. They also report **higher levels of trust** in all datasets where this is measured. Furthermore, they report **stronger ties with their local community**. Another outcome confirmed by every single dataset is the **higher share of volunteers** among members of sport organisations. The only aspect where **there is no consistent difference** between members and non-members is **mixing with people of different backgrounds** (age, race, religion, education).

However, the sports club member and non-member subsamples are also quite different demographically. Males are the most prominent over-represented category by far among sports club members. Furthermore, sports club members are also generally younger, richer, have a lower share of ethnic minorities, come from less deprived areas and higher socio-economic groups, and are more educated. All these factors are likely to drive differences in wellbeing and social outcomes instead of sport club membership, and we therefore control for them in a regression to take the analysis further.

3.2.2. Regression analysis

Descriptive statistics such as the above can only show us how a social outcome or demographic statistic varies between sports club members and non-members in isolation. In order to find out whether the differences in the outcomes can be ascribed to sports club membership or to other factors, we move on to regression analysis, the results of which are presented in Table 6 below. (An example of full regression output including the coefficients of all control variables is available in Annex 1.)

Different regression specifications have been tried, but it was ultimately settled that the panel component of the datasets in question is insufficient for meaningful analysis using fixed effects modelling. The Taking Part and Community Life datasets only have ca. 10-15% of respondents sampled more than once.

Even the dataset with the most robust panel component – Understanding Society – suffers from the fact that sport group membership and social outcomes are not collected in every wave, and thus even this dataset has an average time dimension of only 1.5 observations per respondent. We therefore fall back to OLS regressions to present the main findings of this study, with FE results for USoc presented for reference.

Table 6. Association between sports club membership and social outcomes: all datasets

Outcome variable	USoc - Age 16 to 25		Taking Part - Age 16 to 25	Community Life - Age 16 to 24	Active Lives - Age 16 to 25	USoc Youth - Age 10 to 15
	OLS	Fixed Effects	OLS	OLS	OLS	OLS
Friends have diverse ages (1 to 4)	0.007	0.017		-0.061		
Friends have diverse ethnicities (1 to 4)	0.027	-0.074*		0.085*		
Friends have diverse faiths (1 to 4)				-0.056		
Friends have diverse education levels (1 to 4)	-0.020	0.047		-0.097		
Friends live in different areas (1 to 4)	-0.075***	0.035				
Has friends (0/1)	0.007***	0.003				
Number of close friends	0.345***	0.300***				0.432***
Can rely on friends (1 to 4)	0.059***	0.008				
Satisfaction with friends (1 to 7)						0.073***
Loneliness (1 to 4)				-0.002		
Trust people in general (1 to 3)				0.121*		
Trust people in neighbourhood (1 to 5)	0.035*	0.027		0.102**	0.075**	
Trust in institutions: Parliament (1 to 4)				-0.028		
Trust in institutions: the police (1 to 4)				0.137		
Trust in institutions: local council (1 to 4)				0.055		
Talk to people in neighbourh. (1 to 5)	0.133***	0.110*		0.051		
Feel belonging to neighbourh. (1 to 5)	0.075***	0.009				
Feel belonging to local area (1 to 4)			0.071	-0.310**		
People from different backgrounds get on well in local area (1 to 4)			0.050	0.025		
Can influence local area (1 to 4)				0.021		
Satisfaction with local area (1 to 5)				0.125***		
Local friendships mean a lot (1 to 5)	0.090***	0.009				
Volunteered in the last 12 months (0/1)	0.038***	0.044	0.111***	0.077***	0.344***	0.084***
Formal volunteering in the last 12 months (0/1)				0.140***		
Informal volunteering in the last 12 months (0/1)				0.031		
Hours volunteered in the last 4 weeks	0.106	4.619				
Social mixing during formal volunteering (0/1)				0.135***		
Spend time with family (0/1)			0.030**			0.047**
Frequency of meeting family and friends (1 to 4)				0.068		

Can achieve own goals (1 to 5)					0.122***	
Life satisfaction (1 to 7 or 0 to 10)	0.074***	0.028	0.192**	0.114	0.342***	0.111***
Happiness (0 to 10)			0.231***	0.305***	0.231***	
Anxiety (0 to 10)			0.112	-0.145	-0.038	
Worthwhile life (0 to 10)			0.209**	0.129	0.389***	
General health (1 to 5)	0.270***	0.068***	0.161***	0.085**		0.286***
Mental health problems - GHQ index, 0(best) to 36(worst)	-0.514***	-0.069				
Body Mass Index					-0.304**	

Notes: Model, sample and dataset specified in column header. Coefficients of 'sports club membership' variable shown. All models include control variables for a wide range of determinants of health and wellbeing as set out in Fujiwara and Campbell (2011). List of control variables provided in Section 2.2.4. Stars indicate statistical significance levels: *** < 1%; ** < 5%; * < 10% significance. Heteroskedasticity-robust standard errors used.

Consistently across different datasets we can see a positive association for respondents aged 25 or below between sports club membership and the following outcomes (compared to those who are not members of sport clubs):

- **Social connections** – having friends, number of close friends, relying on friends, satisfaction with friends
- **Trust** – trusting people in general, trusting neighbours. No significant relationship for trust in institutions
- **Community cohesion** – talking to people in the local area, belonging to local area, satisfaction with local area
- **Volunteering** – consistently across all datasets; in particular formal volunteering (Community Life) and volunteering in sport (Active Lives)
- **Perceived ability to achieve goals** (perseverance) – in the Active Lives data.
- **Life satisfaction** (all datasets) and **happiness** (wherever measured)
- **Health** (unsurprisingly) – all datasets.

The relationship between sports club membership and social mixing (in terms of having friends from diverse backgrounds) is less clear.

Overall, this paints a picture of the average sports club member as more socially engaged and involved in the life of their community. Keeping in mind the caveats identified in the Methodology section, one can nevertheless argue that participation in sport organisations is a useful means to increase trust and social capital. It therefore emerges as a potential policy recommendation targeting vulnerable lower socio-economic groups, where a deficit in trust and social capital has been identified in Section 1. The next subsection will further support this argument by showing that the association between sport club membership and the social outcomes in question is **stronger for lower socio-economic groups**.

3.2.3. Disaggregated analysis (heterogeneous effects)

A key aspect of this study is to investigate whether the potential benefits of sporting groups membership (supported by evidence in the form of a positive association with various social outcomes) are stronger or weaker for different subpopulations, including vulnerable groups in

particular. For this purpose we ran heterogeneous effects analysis by interacting our main independent variable (sports club membership) with demographic classifiers such as region, rural/urban residence, income-based and occupation-based socio-economic group as well as immigration status (where possible). The results are presented in Tables 7A-7D below.

We can see that most – but not all – of the **wellbeing and social outcomes correlate more strongly with sport club membership for lower socio-economic groups**. This is true for talking to neighbours in Understanding Society, volunteering and the sense of belonging to the local area in Taking Part, trusting neighbours and formal volunteering in Community Life, and perseverance, volunteering and life satisfaction in the Active Lives data.

The implication is that **encouraging sporting group participation would be an effective approach to bridging the gap in trust and community cohesion** between affluent and vulnerable groups, because those that suffer from a trust deficit the most also get the highest benefits (provided that causality can be reasonably established through more robust studies). Sporting groups therefore prove to be a valuable target for both public and private investment, as the resulting welfare benefits to society can be considerable.

Table 7A. Heterogeneous effects analysis in the Understanding Society panel waves 2, 3, 5, 6, as well as Understanding Society Youth waves 2, 4, 6

Dependent variable / Disaggregation criterion	Age 16 to 25				Age 10 to 15	
	Number of close friends	Trust neighbours	Talking to neighbours	Volunteerin g	Number of close friends	Volunteering
Region (grouped)						
North of England	0.389***	0.018	0.130**	0.033	0.484***	0.091***
Midlands	0.333***	-0.003	0.141**	0.042*	0.437***	0.078***
South and East England	0.354***	0.072**	0.174***	0.037*	0.525***	0.095***
London	0.451***	0.032	0.018	0.055*	0.421***	0.041
Wales	0.254*	-0.011	0.034	0.026	0.341**	0.061*
Scotland	0.207*	0.068	0.200**	0.049	0.307**	0.155***
Northern Ireland	0.299*	0.061	0.210**	0.024	0.190	0.048
Urbanisation						
Urban	0.357***	0.032	0.121***	0.034***		
Rural	0.299***	0.047	0.178***	0.055***		
SEG (based on occupation)¹⁹						

¹⁹ We suggest focus on the occupational lower SEG measure. The income-based measure is not available in Active Lives, because there is no income data there.

Higher SEG	0.389***	-0.021	0.125**	0.064***		
Lower SEG	0.354***	0.024	0.162***	0.043***		
Other SEG ²⁰	0.318***	0.064**	0.114***	0.024		
SEG (income proxy)						
Income at or above median	0.339***	0.078***	0.162***	0.037***	0.419***	0.087***
Income below median	0.352***	-0.017	0.099***	0.040***	0.446***	0.080***
Immigration status						
Native	0.345***	0.039*	0.135***	0.037***		
First-generation immigrant	0.338	-0.274	0.169	0.096		
Second-generation immigrant	0.340	0.051	-0.004	0.061		

Notes: OLS regressions. Coefficients of 'sports club membership' interacted with the variable in the row header are shown. All models include control variables for a wide range of determinants of health and wellbeing as set out in Fujiwara and Campbell (2011) and described in Section 3.2.4. Stars indicate statistical significance levels: *** < 1%; ** < 5%; * < 10% significance. Heteroscedasticity-robust standard errors used.

Table 7B. Heterogeneous effects analysis in the Taking Part waves 8-11

Dependent variable / Disaggregation criterion	Age 16 to 25		
	Belong to local area	Spend time with family	Volunteerin g
Region (grouped)			
North of England	-0.050	0.040*	0.118**
Midlands	0.150	0.064**	0.140***
South and East England	0.116	0.016	0.092**
London	0.074	-0.029	0.087
Urbanisation			
Urban	0.048	0.029*	0.104***
Rural	0.189	0.034	0.146**
SEG (based on occupation)			
Other SEG ²¹	0.048	0.044**	0.105***

²⁰ Including students, retired people, and those not in the labour force for other reasons.

²¹ Including students, retired people, and those not in the labour force for other reasons.

Higher SEG	0.046	0.005	0.063
Lower SEG	0.159	0.023	0.180***
SEG (income proxy)			
Income at or above median	0.035	0.047**	0.112***
Income below median	0.092	0.021	0.111***

Notes: OLS regressions. Coefficients of 'sports club membership' interacted with the variable in the row header are shown. All models include control variables for a wide range of determinants of health and wellbeing as set out in Fujiwara and Campbell (2011) and described in Section 3.2.4. Stars indicate statistical significance levels: *** < 1%; ** < 5%; * < 10% significance. Heteroskedasticity-robust standard errors used.

Table 7C. Heterogeneous effects analysis in the Community Life data, waves 1-5

Dependent variable / Disaggregation criterion	Age 16 to 24				
	Trust people in general	Trust neighbours	Volunteerin g	Formal volunteering	Informal volunteering
Region (grouped)					
North of England	0.073	0.165**	0.091**	0.184***	0.057
Midlands	0.242*	0.028	0.047	0.084	-0.030
South and East England	0.122	0.125*	0.093***	0.142***	0.054
London	0.054	0.036	0.056	0.129**	0.015
Urbanisation					
Urban	0.102	0.088*	0.077***	0.135***	0.032
Rural	0.266	0.188*	0.076	0.171***	0.026
SEG (based on occupation)					
Higher SEG	0.073	0.019	0.055	0.039	0.044
Lower SEG	0.033	0.189**	0.090**	0.147***	0.012
Other SEG ²²	0.174*	0.102*	0.081***	0.186***	0.034
SEG (income proxy)					
Income at or above median	0.105	0.101	0.075*	0.159***	0.022
Income below median	0.124*	0.102**	0.077***	0.136***	0.034

²² Including students, retired people, and those not in the labour force for other reasons.

Notes: OLS regressions. Coefficients of 'sports club membership' interacted with the variable in the row header are shown. All models include control variables for a wide range of determinants of health and wellbeing as set out in Fujiwara and Campbell (2011) and described in Section 3.2.4. Stars indicate statistical significance levels: *** < 1%; ** < 5%; * < 10% significance. Heteroskedasticity-robust standard errors used.

Table 7D. Heterogeneous effects analysis in the Active Lives data, wave 2

Dependent variable / Disaggregation criterion	Age 16 to 25				
	Can achieve goals	Trust people in local area	Volunteerin g	Life satisfaction	BMI
Region (grouped)					
North of England	0.116**	0.149**	0.355***	0.308**	0.101
Midlands	0.096	0.030	0.355***	0.247*	-0.081
South and East England	0.123***	0.082	0.358***	0.319***	-0.739***
London	0.174**	-0.003	0.269***	0.618***	-0.093
Urbanisation					
Urban	0.100***	0.057	0.333***	0.328***	-0.339***
Rural	0.261***	0.188**	0.410***	0.432**	-0.094
SEG (based on occupation)					
Higher SEG	0.077	0.079	0.264***	0.176*	-0.232
Lower SEG	0.192***	-0.023	0.435***	0.545***	-0.355
Other SEG ²³	0.131***	0.104**	0.368***	0.388***	-0.337**

Notes: OLS regressions. Coefficients of 'sports club membership' interacted with the variable in the row header are shown. All models include control variables for a wide range of determinants of health and wellbeing as set out in Fujiwara and Campbell (2011) and described in Section 3.2.4. Stars indicate statistical significance levels: *** < 1%; ** < 5%; * < 10% significance. Heteroskedasticity-robust standard errors used.

²³ Including students, retired people, and those not in the labour force for other reasons.

4. Conclusion

We see in this research that the baseline (starting point) levels for the outcomes such as life satisfaction, health, trust and volunteering are consistently lower for lower socio-economic groups. We also observe in the data that participation in sport groups is associated with higher levels of wellbeing, health, trust and social connection.

Furthermore, the uplift associated with participating in a sport club/organisation is often significantly higher for lower socio-economic groups as opposed to higher socio-economic groups. The suggestion is that these groups have significantly more to gain from sports club membership.

This then helps to explain the key reason for the work of Sported to promote participation in sport groups as one way to reduce the trust and wellbeing gap between different groups in society, thus both increasing total welfare and contributing to a more equal welfare distribution.

The following final caveat should be noted. Despite controlling for many important determinants of wellbeing and social capital, we cannot rule out the suggestion that reverse causality is at work. In other words, people who are already social, integrated, and confident in society are more likely to be members of a sports club.

Conversely, those who are less confident, and more socially isolated are less likely to do so. We hope that future data collection efforts will allow the application of more robust methodology (such as Instrumental Variables, Panel Data Regressions or Difference-in-Differences) that will provide more certainty that the positive association identified here actually represents a direct benefit of sport club membership.

Annex 1. Example Regression Results

Table A1. Full OLS regression results and variable descriptions: Understanding Society, respondents aged 16-25. Dependent variable: Number of close friends.

Variable / level description	Regression coefficient
Number of close friends (6-9 as 6, 10-19 as 7, 20+ as 8)	Dependent variable
Sports club membership	0.345***
Log equivalised household income	0.093***
Age	-0.087
Age squared	0.002
Gender: male (reference group = female)	0.147***
Marital status:	
Married or civil partner	-0.309***
Divorced or former civil partner	-0.593
Separated from husband / wife / civil partner	-0.818
Single (reference group)	0.000
Living as couple	-0.257***
Education:	
Degree (reference group)	0.000
Other higher education	-0.544***
A-level or equivalent	-0.518***
GCSE or equivalent	-0.742***
Other qualifications	-1.182***
No qualifications	-1.020***
Employment status:	
Self-employed	0.318**
In paid employment (full or part-time)	0.216***
Unemployed (reference group)	0.000
On maternity leave	-0.033
Looking after family or home	0.130
Full-time student	0.535***
Long-term sick or disabled	0.066
On a government training scheme	0.270
Unpaid worker in family business	-0.565
Doing something else	-0.100
Urban / rural area:	
Urban area (reference group)	0.000
Rural area	0.051
General health (self-perceived):	
Poor (reference group)	0.000
Fair	0.311*
Good	0.352**
Very good	0.440***

Excellent	0.492***
Number of children:	
0 (reference group)	0.000
1	-0.118
2	-0.260**
3	-0.021
4 or more	0.136
Respondent is religious:	
No (reference group)	0.000
Yes	0.165***
unknown	0.163***
Respondent is a carer for someone in the household:	
No (reference group)	0.000
Yes	0.094
Region of residence:	
North east (reference group)	0.000
North west	0.470***
Yorkshire and the humber	0.282**
East midlands	0.311***
West midlands	0.305***
East of England	0.573***
London	0.427***
South east	0.542***
South west	0.429***
Wales	0.538***
Scotland	0.590***
Northern Ireland	0.011
Accommodation ownership:	
Owned outright (reference group)	0.000
Owned/being bought on mortgage	-0.020
Shared ownership (part-owned part-rented)	-0.197
Rented	-0.149***
Rent-free	-0.147
Other	-0.697**
Respondent would prefer to move house:	
No (reference group)	0.000
Yes	-0.015
Broad ethnic category:	
White (reference group)	0.000
Mixed	-0.258**
Asian	-0.567***
Black	-0.644***
Other	-1.099***
Unknown	-0.062

Wave of survey:	
3 – 2011/12 (reference group)	0.000
6 – 2014/15	-0.185***
Interview month:	
January (reference group)	0.000
February	0.129
March	0.055
April	0.151*
May	-0.079
June	-0.039
July	0.093
August	-0.017
September	0.046
October	-0.031
November	0.029
December	-0.022
Constant	3.939***
Observations	10630
Adjusted R-squared²⁴	0.088

²⁴ Note that regressions of wellbeing and social outcomes tend to have low (adjusted) R-squared values, usually 0.1-0.3, even taking into account all the sociodemographic controls. Wellbeing and social outcomes such as trust or having friends are complex phenomena and there is of course a lot of variation left unexplained.