Value of Swimming

Technical Report

Swim England

jump

Contents

Glossary

AL	Active Lives survey
ALY	Active Lives Children and Young People survey
BHPS	British Household Panel survey
BMI	Body Mass Index
DCMS	Department for Digital, Culture, Media and Sport
FAS	Family Affluence Score
FE	Fixed Effects
GCSE	General Certificate of Secondary Education
GHQ	General Health Questionnaire
HE	Higher Education
LSOA	Lower Layer Super Output Areas
NS-SEC	National Statistics Socio-economic Classification
OECD	Organisation for Economic Cooperation and Development
OLS	Ordinary Least Squares regression
ONS	Office for National Statistics
SEC	Socio-economic Class
SEG	Socio-economic Groups
TP	Taking Part survey
USoc	Understanding Society survey
USoc Youth	Understanding Society Youth survey

Introduction

As the national governing body for swimming, water polo, diving and synchronised swimming in England, Swim England is committed to championing the benefits of water-based activity.

Whilst the health and social benefits of sport and physical activity are widely recognised, there is a need to build a robust evidence base around the specific benefits of water-based activity. Swim England wanted to take on this challenge, and show the role the sport has in improving people's lives.

In this research, five large, relevant national datasets are analysed using advanced, robust, statistical methods to reveal the health and wellbeing increases observed in swimmers relative to non-swimmers.

The Government's ***Sporting Future***[1] strategy seeks to encourage greater levels of activity. It has a clear focus on five outcomes:

- Physical wellbeing
- Mental wellbeing
- Individual development
- Social and community development
- Economic development.

Several identified key performance indicators within the strategy align with expected benefits of swimming. These five outcomes, therefore, offer a clear structure on which to understand the different ways that benefit and value is created through regular swimming.

1 https://www.gov.uk/government/publications/sporting-future-a-new-strategy-for-an-active-nation

Key findings

All the key findings below are true for both male and female participants, unless specifically stated.

Adult (16+ years)

1. Swimmers report considerably higher levels of wellbeing and health compared to non-swimmers. This is true for all wellbeing measures investigated:
    a. life satisfaction - on average 1.9 per cent more satisfied with life
    b. general health - feeling on average 6.4 per cent healthier; this association is comparable to feeling up to 12 years younger
    c. happiness - feeling on average 1.1 per cent happier, equivalent to moving from an urban to a rural areas; this association is comparable to moving to a less deprived area
    d. reduced anxiety - feeling on average 4.1 per cent less anxious; this is half of the anxiety reduction felt from retiring
    e. worthwhile life - a 1 per cent increase in feeling that life is more worthwhile; this is the equivalent uplift of having a third child.

2. All the datasets show swimmers to be more socially connected and engaged in their community.
    a. Social connections – swimmers:
        i. can rely on their friends more - by 3.1 per cent
        ii. are 26.7 per cent less likely to have no friends, indicating less loneliness[2].
    b. Volunteering – swimmers:
        i. are 26.1 per cent more likely to volunteer
        ii. are 34.4 per cent more likely to volunteer in sport or other physical activity
        iii. who do volunteer spend about 7.5per cent more hours giving unpaid help.
    c. Social trust – swimmers:
        i. have 2.7 per cent more trust in people in general and 3.6 per cent more trust in their neighbours.

3. The positive association between health and swimming increases considerably as people get older (Table 7G).

4. Compared to men, women generally have a stronger association between swimming and wellbeing outcomes, being able to achieve goals, and rely on friends more (Table 7A). In particular for both women and girls, the impact on confidence in achieving your own goals among swimmers can be double or more for women over men.

5. More frequent swimming is associated with a higher wellbeing uplift, with swimming once to twice a week having around double the positive wellbeing correlation of swimming once a month (Table 6B).

6. Swimmers from higher SEG experience stronger associations between swimming and life satisfaction than swimmers from lower SEG. Swimmers from lower SEG have a stronger correlation with volunteering (Table 7A).

7. Outdoor swimming seems to generate twice the happiness of swimming indoors (Table 6A).

2 This is not the same as being 26.7 per cent more likely to have friends, because the likelihood to have friends is already relatively close to 1.

Table 1A. Key associations between swimming and DCMS outcomes - adults

DCMS outcome	Indicator variable	OLS regression coefficient	Dataset	Interpretation
Physical wellbeing	General health (1 to 5 scale)	0.153***	USoc	Swimming is associated with being about 6.4% healthier compared to the average population.
Physical wellbeing	Mental health (0 to 36 scale - GHQ-12 index[3])	-0.353***	USoc	Swimming is associated with having around 3.1% better mental health compared to the average population.
Physical wellbeing	Body Mass Index	-0.364***	AL	Swimming is associated with having ca. 1.4% lower body mass compared to the average population.
Mental wellbeing	Life satisfaction (1 to 7)	0.080***	USoc	Swimming is related to being about 1.9% more satisfied with life compared to the average population.
Mental wellbeing	Happiness (0 to 10)	0.084***	TP	Swimming is connected to being about 1.1% happier compared to the average population.
Mental wellbeing	Anxiety (0 to 10)	-0.133***	AL	Swimming is associated with being about 4.1% less anxious than the average population.
Mental wellbeing	Sense of worth-while life (0 to 10)	0.077***	TP	Swimming is associated with feeling 1% more that your life is worthwhile.
Individual development	Can achieve own goals (1 to 5)	0.068***	AL	Swimming is associated with 2.4% more confidence that the respondent can achieve the goals he/she sets him/herself.
Social and community development	Having friends (0/1)	0.012***	USoc	Swimming is connected with 26.7% lower likelihood compared to the average population of the respondent having no friends.
Social and community development	Can rely on friends (1 to 4)	0.068***	USoc	Swimming is related to relying on one's friends about 3.1% more compared to the average level in the population.
Social and community development	Trust people in general (1 to 3)	0.025*	TP	Swimming is associated with higher levels of trust in people - the difference is about 2.7% of the average levels of trust.
Social and community development	Trust people in neigh-bourhood (1 to 5)	0.076***	TP	Swimming is associated with 3.6% higher trust in one's neighbours.
Social and community development	Volunteered in the last 12 months (0/1)	0.047***	USoc	Swimming is associated with an increase in the likelihood to volunteer which is around 26% of the average volunteering rate among UK adults.
Social and community development	Volunteered in sport in the last 12 months (0/1)	0.074***	AL	Swimming is likewise associated with an increase in volunteering in sports of about 34% of the average.
Social and community development	Hours volunteered in the last 4 weeks	0.149	USoc	Swimming is associated with volunteering 7.4% more hours (compared to the average hours volunteered).

*Notes: OLS regressions. Coefficients of swimming 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 swimmers stand as opposed to non-swimmers. 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 swimmers stand x% closer to the positive end of the (e.g.) trust spectrum than non-swimmers. 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 Annex 1. Stars indicate statistical significance levels: *** < 1%; ** < 5%; * < 10% significance. Heteroskedasticity-robust standard errors used. Young people (7-16 years).*

3 https://www.gl-assessment.co.uk/products/general-health-questionnaire-ghq/

Young people (7-16 years)

8. Just like adults, young swimmers have higher wellbeing than non-swimmers and are happier, and healthier. They also show higher levels of self-confidence and self-efficacy.

a. General health - feeling on average 4.6 per cent healthier.
b. Life satisfaction - on average 1.2 per cent more satisfied with life.
c. Happiness - on average 4.1 per cent happier.
d. Worthwhile life - 4.8 per cent higher feeling that life is more worthwhile.
e. Self-confidence - on average 3.3 per cent more confident that they can achieve their own goals.

9. The data shows young swimmers to be more socially connected and engaged in their community. They are also more satisfied with their friendships, spend more time with their families and volunteer more.

a. Social connections – young swimmers:
   i. are 2.4 per cent more likely to eat evening meals with their families
   ii. have 0.15 extra close friends (equivalent to 1 in 7 young swimmers youth respondents having an extra close friend), which represents a 3.2 per cent increase .
b. Volunteering – young swimmers:
   i. are 17.1 per cent more likely to volunteer in sport or other physical activity
   ii. volunteer about 19.2 per cent more often[4] .
c. Social trust – young swimmers:
   i. have 2.1 per cent more trust in people in their neighbourhood.

10. Girls who swim have considerably higher increases in wellbeing, health and self-confidence compared to boys.

Table 1B. Key associations between swimming and DCMS outcomes - youth

DCMS outcome	Indicator variable	OLS regression coefficient	Dataset	Interpretation
Physical wellbeing	General health (1 to 5 scale)	0.130***	USoc Youth	Swimming is associated with being about 4.6% healthier compared to the average young population.
Mental wellbeing	Happiness with life overall (1 to 7)	0.061***	USoc Youth	Swimming is related to being about 1.2% more satisfied with life compared to the average young population.
Mental wellbeing	Happiness (0 to 10)	0.294***	ALY	Swimming is connected to being about 4.1% happier compared to the average young population.
Mental wellbeing	Sense of worthwhile life (0 to 10)	0.323***	ALY	Swimming is associated with feeling 4.8% more that your life is worthwhile.
Individual development	Can achieve own goals (1 to 4)	0.074***	ALY	Swimming is associated with 2.4% more confidence that the respondent can achieve the goals he/she sets him/herself.
Social and community development	Number of close friends	0.150***	USoc Youth	Swimming is associated with having 0.15 more close friends (equivalent to 1 in 7 youth respondents having an extra close friend), which is 3.2% of the average population level.
Social and community development	Spend time with family (1 to 4)	0.076***	USoc Youth	Young respondents who swim are (all other things equal) more likely to eat evening meals with their families by an amount equal to 2.4% of the average level.
Social and community development	Trust people in neighbourhood (1 to 4)	0.042***	ALY	Young swimmers are about 2.1% more likely to believe most neighbours can be trusted rather than none of the neighbours (after controlling for other factors).

4 The value is approximate because volunteering frequency is grouped into categories in the data.

Social and community development	Volunteered in sport in the last 12 months (0/1)	0.065***	ALY	Swimming is associated with an increase in the likelihood to volunteer in sport which is around 17% of the average volunteering rate among UK youth.
Social and community development	Frequency of volunteering (1 to 6)	0.244***	USoc Youth	Swimming is associated with approximately a 19% increase in the frequency of volunteering.

Notes: Same as Table 1A above.

Results in full

In this section we present the results of the analysis.

The descriptive statistics presented in Tables 2A-2E show how swimmers generally have higher levels of wellbeing, trust and social capital than non-swimmers. This finding is taken further by the regression analysis results (Table 3), which show that the positive association between swimming 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 swimming.

Finally, disaggregated analysis shows patterns that vary between datasets, although some trends are more consistent. Life satisfaction is consistently more strongly associated with swimming for higher SEG, whereas lower SEG have a stronger correlation with volunteering. Also, older respondents have the highest health uplift associated with swimming. Women experience higher increases in life satisfaction and self-confidence associated with swimming – this is especially true for younger girls.

Descriptive statistics

The tables below compare the main wellbeing and social outcomes, as well as the demographic composition of swimmers and non-swimmers in five datasets:

- Understanding Society (UK-wide)
- Understanding Society Youth (UK-wide)
- Active Lives (England)
- Active Lives Children and Young People (England)
- Taking Part (England).

Understanding Society (USoc)

In the USoc data we can see that swimmers are happier and healthier – they report having better general health as well as mental health. They are less likely to be lonely (have no friends at all), and they are also able to rely upon their friends more. Furthermore, they are about 50 per cent more likely to volunteer, and those who do volunteer spend about 10 per cent more hours giving unpaid help.

Demographically swimmers are unsurprisingly considerably younger than non-swimmers (10 years younger on average), and more likely to be female and have children. A key aspect we can clearly see is the higher socio-economic background that swimmers come from. Their household income is more than 25 per cent higher, they are more than 50 per cent more likely to have a university degree, more likely to be employed and less likely to be part of an ethnic minority.

Table 2A. Wellbeing, social outcomes and demographics by swimming status in the USoc panel dataset, waves 2 and 5

Swimming or diving in the last 12 months	No	Yes	Total
Sample size	60907	29287	90194
Wellbeing outcomes			
Life satisfaction (1 to 7)	5.06	5.29	5.14
General health (1 to 5)	3.23	3.71	3.38
Mental health problems - GHQ index, 0(best) to 36(worst)	11.42	10.79	11.21
Individual development, and social and community development outcomes			
It is easy for me to stick to my aims and accomplish my goals (1 to 4), mean	2.97	3.03	2.99
Has no friends	5.8% (3040/52713)	2.1% (572/26910)	4.5% (3612/79623)
I can rely upon my friends (1 to 4), mean	3.17	3.29	3.21
Volunteered in the last 12 months	15.6% (5269/33826)	23.1% (3804/16446)	18.0% (9073/50272)
Hours volunteered in the last 4 weeks, mean	1.9	2.21	2
Demographics			
Household income[5] (monthly)	£3,237	£4,170	£3,540
Age	50.96	40.98	47.72
Female	54.2% (33026/60907)	59.2% (17324/29287)	55.8% (50350/90194)
Married	52.1% (31719/60855)	51.9% (15176/29261)	52.0% (46895/90116)
No children	77.0% (46881/60907)	62.1% (18179/29287)	72.1% (65060/90194)
Higher education degree	18.6% (11240/60513)	30.2% (8777/29021)	22.4% (20017/89534)
Employed (full or part-time)	40.7% (24799/60880)	58.6% (17142/29275)	46.5% (41941/90155)
Urban area	75.2% (45745/60821)	75.2% (21979/29238)	75.2% (67724/90059)
Religious	59.0% (30147/51139)	50.2% (11837/23580)	56.2% (41984/74719)
Lives in Greater London	12.5% (7632/60821)	10.1% (2963/29238)	11.8% (10595/90059)
White	82.4% (50204/60907)	88.9% (26048/29287)	84.5% (76252/90194)
Current job socio-economic classification is low[6]	29.9% (18188/60768)	36.4% (10633/29183)	32.0% (28821/89951)
Equivalised household income is below median	56.7% (34512/60899)	39.6% (11599/29285)	51.1% (46111/90184)
Immigration status: native	98.5% (59969/60907)	98.9% (28952/29287)	98.6% (88921/90194)

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

5 Contains the earnings of everyone living in that household (including parents if respondent lives with his/her parents) from all sources of income.
6 Note that people not in the labour force (full-time students, retired people and others) are outside the NS-SEC classification, making up the so-called NS-SEC residuals.

Understanding Society Youth (USoc Youth) (aged 10-15)

Children who swim or dive have higher wellbeing and social capital – they are happier, healthier, more satisfied with their friendships, spend more time with their families and volunteer more. They also agree more with all statements indicating self-confidence except for "being able to solve one's own problems." Correspondingly, they are less likely to agree with all statements indicating a lack of self-confidence.

In terms of their socio-economic backgrounds, the situation is rather similar to that of adults in the USoc data, though the differences are not so pronounced. Swimmers aged 10 to 15 tend to come from slightly richer families, are slightly more female, about one year younger on average and are a little more likely to come from rural areas. They have more siblings but fewer adults in the household.

Table 2B. Wellbeing, social outcomes and demographics by swimming status in the USoc Youth panel dataset, waves 2, 4, 6

	Age 10-15		
Does swimming or diving (timeframe not specified)	No	Yes	Total
Sample Size	7644	4676	12320
Wellbeing outcomes			
Life satisfaction (1 to 7)	5.84	6.03	5.91
General health (1 to 5)	3.74	3.94	3.81
Social and community development outcomes			
Number of close friends, binned (1 to 8), mean	4.61	4.67	4.63
Satisfaction with friends (1 to 7), mean	6.27	6.35	6.3
Frequency of eating evening meals with family (1 to 4), mean	3.1	3.24	3.15
Volunteers more than once a year	33.0% (2491/7542)	41.3% (1904/4605)	36.2% (4395/12147)
Self-confidence outcomes - % agree or strongly agree			
I feel I have a number of good qualities	94.6% (7149/7559)	96.9% (4478/4622)	95.5% (11627/12181)
I don't have much to be proud of	20.1% (1519/7567)	15.1% (698/4637)	18.2% (2217/12204)
I certainly feel useless at times	39.8% (3009/7564)	36.8% (1702/4623)	38.7% (4711/12187)
I am as able as most people	90.7% (6860/7564)	91.3% (4233/4634)	90.9% (11093/12198)
I am a likeable person	95.1% (7177/7550)	96.4% (4464/4629)	95.6% (11641/12179)
I can usually solve my own problems	89.0% (6729/7559)	88.2% (4075/4621)	88.7% (10804/12180)
I am inclined to feel I am a failure	9.9% (743/7474)	8.5% (389/4567)	9.4% (1132/12041)
At times I feel I am no good at all	27.8% (2099/7550)	24.6% (1140/4629)	26.6% (3239/12179)
Demographics			
Household income (monthly)	£3,953	£4,121	£4,017
Age	12.91	11.95	12.55
Number of adults aged 14+ in household (OECD definition)	2.94	2.6	2.81
Number of children aged 0-13 in household, mean	1.52	1.77	1.61
Female	48.9% (3737/7644)	50.9% (2380/4676)	49.7% (6117/12320)
Urban area	77.6% (5924/7634)	74.4% (3474/4668)	76.4% (9398/12302)
Lives in Greater London	13.1% (999/7634)	13.2% (615/4668)	13.1% (1614/12302)
Equivalised household income is below median	51.8% (3945/7619)	45.5% (2120/4664)	49.4% (6065/12283)

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

Active Lives (AL)

The AL dataset introduces three other ONS wellbeing measures - happiness and a sense of a worthwhile life, which are considerably higher for swimmers along with life satisfaction, and the inverted wellbeing measure of anxiety, which is accordingly lower.

There is no general health indicator in the AL data, but it can be proxied by the BMI, which is lower for swimmers, who also have a substantially lower share of obese respondents.

Furthermore, swimmers score higher on both the individual and social and community development indicators - perseverance and trust - and also volunteer in sport much more frequently. Confirming the demographic trend first unveiled in the USoc data, they are younger, more likely to be female, living with a partner, and with children. They also come from more affluent areas, and have slightly higher shares of rural and white respondents.

Note - for a finer breakdown of the association between swim frequency and wellbeing, see table 6B.

Table 2C. Wellbeing, social outcomes and demographics by swimming status in the AL dataset, wave 2

Swimming in the last 12 months (incl. diving, water polo and water-based class)	No	Yes	Total Observations
Sample size	73089	34380	107469
Wellbeing outcomes			
Life satisfaction, 0-to-10 scale, mean	7.05	7.35	7.15
Happiness, 0-to-10 scale, mean	7.13	7.36	7.2
Anxiety, 0-to-10 scale, mean	3.25	3.15	3.22
Worthwhile life, 0-to-10 scale, mean	7.34	7.61	7.42
Respondent's BMI, mean	26.13	25.26	25.85
Respondent has "normal weight" according to BMI	45.0% (28067/62302)	53.8% (16069/29892)	47.9% (44136/92194)
Individual development, and social and community development outcomes			
I can achieve most of the goals I set myself (% agree or strongly agree)	69.1% (24635/35659)	77.5% (12613/16282)	71.7% (37248/51941)
Most of the people in my local area can be trusted (% agree or strongly agree)	47.4% (16875/35615)	54.5% (8871/16265)	49.6% (25746/51880)
Volunteered in the last 12 months to support sport and physical activity	18.1% (6707/37121)	28.7% (5109/17794)	21.5% (11816/54915)
Volunteered in sport more than once in the last 12 months, excluding raising funds	13.2% (4905/37121)	21.4% (3804/17794)	15.9% (8709/54915)
Volunteered in sport in the last 4 weeks, excluding raising funds	8.4% (3107/37121)	12.5% (2230/17794)	9.7% (5337/54915)
Demographics			
Age	49.92	43.26	47.79
Female	51.8% (37845/73049)	59.4% (20426/34367)	54.2% (58271/107416)
Lives as a couple	63.1% (43568/69030)	67.8% (21516/31752)	64.6% (65084/100782)
No children	74.8% (54453/72822)	64.5% (22106/34263)	71.5% (76559/107085)
Level 4 education (Certificate of HE) or above	50.2% (35787/71285)	67.8% (22227/32772)	55.8% (58014/104057)
Employed (full or part-time)	55.7% (39780/71441)	69.1% (22661/32811)	59.9% (62441/104252)
Urban area	80.1% (58578/73087)	78.9% (27133/34380)	79.8% (85711/107467)
Religious	68.8% (24181/35146)	62.7% (10695/17053)	66.8% (34876/52199)
White	89.0% (62387/70126)	92.9% (30939/33317)	90.2% (93326/103443)
LSOA Index of Multiple Deprivation decile, mean	5.4	5.85	5.54
Current job SOC is low[7]	19.4% (14146/73089)	9.7% (3338/34380)	16.3% (17484/107469)

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

7 Note that respondents not in the labour force make up the so-called NS-SEC residuals.

Active Lives Children and Young People (ALY) (aged 5-16[8])

When we look at the youth version of the AL data, we see that the higher wellbeing of swimmers is maintained – they are more satisfied with life overall, happier in the present moment, and more likely to think that their life is worthwhile. Furthermore, they have slightly higher proportions that are perseverant (94 per cent vs. 89 per cent) and trusting (84 per cent vs. 80 per cent), and a have a considerably higher incidence of volunteering in sport.

Furthermore, the demographic differences are rather similar to those unveiled by Usoc Youth – child swimmers are younger, slightly more likely to be female, and come from somewhat more affluent families.

Table 2D. Wellbeing, social outcomes and demographics by swimming status in the ALY dataset, wave 1

Swimming in the last 7 days[9]	No	Yes	Total Observations
Sample size	84175	24150	108325
Wellbeing outcomes			
Life satisfaction, 0-to-10 scale, mean	6.53 (44579)	7.01 (6410)	6.59 (50989)
Happiness, 0-to-10 scale, mean	7.01 (80376)	7.69 (20869)	7.15 (101245)
Worthwhile life, 0-to-10 scale, mean	6.67 (44455)	7.20 (6397)	6.73 (50852)
Individual development, and social and community development outcomes			
Self-efficacy: If I find something difficult, I keep trying until I can do it (% agree or strongly agree)	89.2% (66077/74078)	94.1% (18365/19516)	90.2% (84442/93594)
Trust: How much do you feel you can trust people who are a similar age to you? (% agree or strongly agree)	80.0% (63793/79695)	83.7% (17348/20722)	80.8% (81141/100417)
Volunteered in the last 12 months to support sport and physical activity	53.0% (33869/63892)	64.4% (8023/12461)	54.9% (41892/76353)
Volunteered in sport more than once in the last 12 months, excluding raising funds	36.6% (23008/62843)	47.9% (5788/12091)	38.4% (28796/74934)
Demographics			
Age	11.32 (81720)	9.60 (23619)	10.94 (105339)
Female	50.5% (39565/78317)	54.0% (12327/22810)	51.3% (51892/101127)
Urban area	81.8% (68830/84175)	80.1% (19345/24150)	81.4% (88175/108325)
Has a disability	23.3% (18909/81301)	27.5% (6458/23477)	24.2% (25367/104778)
White British	71.2% (58045/81524)	72.5% (17085/23558)	71.5% (75130/105082)
LSOA Index of Multiple Deprivation decile, mean	6.04 (84175)	6.44 (24150)	6.13 (108325)
Low family affluence[10]	20.7% (16607/80092)	15.7% (3654/23208)	19.6% (20261/103300)
Receipt of free school meals	22.4% (8900/39676)	19.5% (1126/5772)	22.1% (10026/45448)

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

8 Only respondents from year 3 and above (age 7) were asked about relevant outcomes, whereas life satisfaction is only available from school year 7 (age 11).

9 Unfortunately there is no data on swimming in the last 12 months in this dataset. The closest available information was used.

10 Active Lives children has a variable called "Family Affluence Score", which is from 0 to 20 and derived from the respondent's answers to a series of questions about the possessions they have at home and their household expenses.

Taking Part (TP)

The TP dataset confirms the differences in wellbeing outcomes revealed in the AL data (except for anxiety) and the health differential found in USoc, as well as the demographic characteristics mentioned earlier (younger, more female, married, with children, richer).

Swimmers also fare better in terms of social outcomes. In this dataset we can see that they are slightly more trusting and more sociable (spend more time with their friends and family), and volunteer more (confirming the trend in the other data). There is no notable difference in terms of diversity mixing or community belonging.

The TP dataset is particularly useful to us because it distinguishes between indoor and outdoor swimming. The differences between the two types are not that strong, though we can see that most wellbeing indicators are even higher for outdoor swimmers. A particularly notable difference is in the levels of trust. Interesting demographics of outdoor swimmers are the higher age than for indoor swimmers, a gender ratio that is close to the sample mean, and even higher income.

Table 2E. Wellbeing, social outcomes and demographics by swimming status in the TP dataset, waves 1-3, 6-12

Swimming or diving in the last 12 months	No swimming	Any swimming	Indoor Swimming	Outdoor Swimming	Total Observations
Sample size	98890	51944	41650	19040	150834
Wellbeing outcomes					
Life satisfaction, 0-to-10 scale, mean	7.63	7.9	7.87	7.95	7.71
Happiness, 0-to-10 scale, mean	7.67	7.91	7.9	8.03	7.76
Anxiety, 0-to-10 scale, mean	2.9	2.9	2.95	2.85	2.9
Worthwhile life, 0-to-10 scale, mean	7.87	8.13	8.14	8.15	7.95
General health, 1-to-5, self-reported, mean	3.79	4.21	4.2	4.26	3.93
Social and community development outcomes					
Trusts people in general	41.0% (8818/21502)	44.6% (5206/11678)	44.0% (4279/9715)	46.6% (1953/4189)	42.3% (14024/33180)
Trusts most neighbours	39.4% (11545/29301)	43.4% (6903/15910)	42.6% (5627/13218)	47.0% (2692/5725)	40.8% (18448/45211)
Meets friends at least once a week	70.8% (20404/28836)	78.2% (12193/15588)	78.7% (10218/12980)	76.8% (4267/5559)	73.4% (32597/44424)
Feels that he/she belongs to local area (% agree or strongly agree)	76.0% (16184/21303)	74.2% (8155/10988)	74.2% (6151/8291)	74.2% (3010/4059)	75.4% (24339/32291)
People from different backgrounds get on well together (% agree or strongly agree)	84.6% (16297/19274)	84.7% (8622/10183)	84.1% (6461/7679)	85.4% (3215/3763)	84.6% (24919/29457)
Spends time with family and friends	79.5% (67699/85114)	89.8% (40137/44709)	89.5% (31841/35574)	91.3% (15099/16542)	83.1% (107836/129823)
Volunteered in the last 12 months	21.2% (20906/98827)	30.3% (15752/51907)	30.4% (12654/41619)	31.7% (6036/19028)	24.3% (36658/150734)
Volunteered in sport (last 12 months)	2.9% (2829/98890)	6.5% (3389/51944)	6.6% (2737/41650)	7.4% (1411/19040)	4.1% (6218/150834)
Demographics					
Personal earnings (yearly)	£16,902	£22,176	£21,214	£24,773	£18,811

Age	54.97	43.56	42.55	45.11	51.04
Female	53.4% (52837/98890)	60.2% (31249/51944)	62.1% (25861/41650)	54.7% (10423/19040)	55.7% (84086/150834)
Married	47.8% (47103/98476)	55.5% (28700/51758)	54.3% (22548/41503)	59.5% (11299/18976)	50.5% (75803/150234)
No children	78.2% (77286/98817)	57.7% (29955/51913)	54.5% (22690/41625)	63.4% (12065/19036)	71.1% (107241/150730)
Higher education degree (including professional)	18.2% (18000/98890)	32.3% (16799/51944)	32.1% (13389/41650)	34.9% (6647/19040)	23.1% (34799/150834)
Doing paid work (includes self-employed)	43.3% (42842/98883)	67.4% (35034/51944)	67.4% (28079/41650)	70.4% (13402/19040)	51.6% (77876/150827)
Urban area	80.7% (66557/82435)	78.9% (33665/42672)	79.3% (26933/33972)	77.2% (12091/15655)	80.1% (100222/125107)
Religious	76.8% (72450/94351)	70.3% (34861/49621)	70.2% (27891/39732)	70.6% (12845/18185)	74.5% (107311/143972)
White	87.7% (86569/98746)	92.4% (47933/51882)	92.0% (38279/41601)	95.2% (18089/19009)	89.3% (134502/150628)
LSOA Index of Multiple Deprivation decile, mean	5.36	5.94	5.87	6.24	5.55
Low SEC (NS-SEC 5-8)[11]	45.8% (45263/98890)	29.5% (15319/51944)	30.0% (12483/41650)	25.4% (4832/19040)	40.2% (60582/150834)
Personal income is below median[12]	42.9% (34477/80311)	31.4% (14294/45584)	33.0% (12112/36662)	26.7% (4462/16736)	38.7% (48771/125895)

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

Descriptive statistics summary

Swimmers report considerably higher levels of wellbeing and health than non-swimmers. This is true for almost all available wellbeing measures - life satisfaction and general health, mental health (USoc), BMI/obesity (AL), happiness and a sense of worthwhile life (TP, AL, ALY), anxiety (AL).

In terms of social outcomes, a pervasive finding throughout all the datasets is that swimmers are more socially connected - they report having more close friends, relying on their friends more (USoc), and spending more time with their friends and family (TP). They also report higher levels of trust (TP, AL, ALY). Another outcome confirmed by all datasets is the higher share of volunteers among swimmers. The data analysed did not enable us to spot a difference between swimmers and non-swimmers in affinity with the local community and mixing with people of different backgrounds.

However, the swimmer and non-swimmer subsamples are also quite different demographically. There is usually a higher share of females among swimmers than non-swimmers (except if we consider outdoor swimming only). Furthermore, swimmers are 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 observed differences in wellbeing and social outcomes between swimmers and non-swimmers, and we therefore control for them in a regression to take the analysis further.

11 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.

12 Because income is only available as a categorical variable (income brackets), a lot of the respondents are exactly at the median. These are not counted as "below median".

Regression analysis

Descriptive statistics can highlight differences in wellbeing and social outcomes between swimmers and non-swimmers, but cannot tell us whether these differences are due to swimming or due to the influence of some other characteristics that are fundamentally different for swimmers. In order to find out whether the differences in the wellbeing and social outcomes can be ascribed to different demographics, we move on to regression analysis.

Ordinary Least Squares (OLS) regressions

The results of (basic) OLS regressions are presented in Table 3 below.

Table 3. OLS results - Association between swimming and wellbeing/social outcomes

Outcome variable	USoc	TP	AL	USoc Youth - age 10 to 15	ALY - age 7 to 16
Has friends (0/1)	0.012***				
Number of close friends				0.150***	
Can rely on friends (1 to 4)	0.068***				
Satisfaction with friends (1 to 7)				0.021	
Spend time with family (0/1)		0.056***		0.076***	
Frequency of meeting friends (1 to 5)		0.126***			
Trust people in general (1 to 3)		0.025*			
Trust people in neighbourhood (1 to 5)		0.076***	0.075***		0.042***
Feel belonging to local area (1 to 4)		0.025**			
People from different backgrounds get on well in local area (1 to 4)		0.011			
Volunteered in the last 12 months (0/1)	0.047***	0.057***		0.068***	
Hours volunteered in the last 4 weeks	0.149				
Frequency of volunteering, categorical (1 to 6)				0.244***	
Volunteered in sport in the last 12 months (0/1)		0.020***	0.074***		0.065***
Volunteered in sport more than once in the last 12 months (0/1)			0.058***		0.060***
Volunteered in sport in the last 4 weeks (0/1)			0.029***		
Can achieve own goals (1 to 4 / 1 to 5)	0.020***		0.068***		0.074***
Life satisfaction (1 to 7 or 0 to 10)[13]	0.080***	0.071***	0.188***	0.061***	0.300***
Happiness (0 to 10)		0.084***	0.164***		0.294***
Anxiety (0 to 10)		0.049	-0.133***		
Worthwhile life (0 to 10)		0.077***	0.168***		0.323***
General health (1 to 5)	0.153***	0.123***		0.130***	
Mental health problems - GHQ index, 0(best) to 36(worst)	-0.353***				
BMI			-0.364***		

*Notes: Model, sample and dataset specified in column header. Coefficients of 'swimming and diving in the last 12 months' 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 Methodology Section. Stars indicate statistical significance levels: *** < 1%; ** < 5%; * < 10% significance. Heteroscedasticity-robust standard errors used.*

Consistently across different datasets we can see a positive association between swimming and the following outcomes:

- social connections - having friends, number of close friends, relying on friends, spending time with friends and family
- trust - trusting people in general, trusting neighbours
- community cohesion - belonging to local area
- volunteering - in general and in sport (all datasets) also with frequency of volunteering

13 USoc (adult and youth) - 1 to 7 scale. All other datasets - 0 to 10 scale.

- perceived ability to achieve goals (USoc, AL)
- life satisfaction (all datasets) and happiness (wherever measured), worthwhile life (wherever measured) and anxiety (AL, but not TP)
- health (all datasets), mental health (USoc) and BMI (AL)
- for young people we can see positive co-movement with wellbeing, health, self-efficacy, trust, volunteering and social connections (USoc Youth, ALY).

Fixed Effects (FE) regressions

FE regressions[14] offer the advantage of controlling not just for the observable variables included in the regression, but also for unobserved individual-specific traits on condition that they do not vary over time. This allows for a greater degree of robustness of the results by eliminating potential sources of omitted variable bias (making the results closer to indicating a causal effect).

FE estimation requires a panel data structure - individuals must be observed at multiple points in time. However, the panel component of the datasets in question does not allow for meaningful FE modelling.

We nonetheless present the results of FE regressions in Table 4 below, but with the caveat that the low statistical significance of the coefficients is due to limitations of the panel structure of the datasets. Therefore, the degree to which the results below can be used as a robustness check on the main OLS findings is somewhat limited.

Table 4. FE results - Association between swimming and wellbeing/social outcomes

Outcome variable	USoc	TP	USoc Youth - Age 10 to 15
	FE	FE	FE
Has friends (0/1)	0.003		
Number of close friends			0.076
Can rely on friends (1 to 4)	0.012		
Satisfaction with friends (1 to 7)			-0.017
Trust people in general (1 to 3)		-0.059	
Trust people in neighbourhood (1 to 5)		0.071**	
Volunteered in the last 12 months (0/1)		0.049***	0.023
Frequency of volunteering, categorical (1 to 6)			0.094*
Volunteered in sport in the last 12 months (0/1)		0.017***	
Spend time with family (0/1)		0.040***	0.059**
Frequency of meeting friends (1 to 5)		0.129***	
Life satisfaction (1 to 7 or 0 to 10)	0.031	-0.032	0.044
Happiness (0 to 10)		0.082***	
Anxiety (0 to 10)		-0.010	
Worthwhile life (0 to 10)		-0.007	
General health (1 to 5)	0.006	0.097***	0.022
Mental health problems - GHQ index, 0(best) to 36(worst)	0.008		

*Notes: Model, sample and dataset specified in column header. Coefficients of 'swimming and diving in the last 12 months' 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 Methodology Section. Stars indicate statistical significance levels: *** < 1%; ** < 5%; * < 10% significance. Heteroscedasticity-robust standard errors used.*

The TP dataset has the advantage of having the biggest number of waves (12, of which most contain the relevant variables). This allows us to observe a significant association of swimming with volunteering, spending time with family/friends, trust in neighbours, happiness and health. Unfortunately, data on life satisfaction and the other wellbeing outcomes only began to be collected recently in the TP data - this is the most likely explanation for the insignificant association. Self-confidence / self-efficacy outcomes for children aged 10-15.

14 For more information regarding methodology, see Annex 1 (Detailed Methodology).

Self-confidence / self-efficacy outcomes for children aged 10-15

The USoc Youth dataset collects a range of indicators of the respondents' self-confidence. These indicators consist of (dis)agreement with specific statements. Some statements are worded in such a way that agreement indicates confidence and disagreement indicates a lack of confidence, whereas other statements function in the opposite manner.

The association between swimming and agreement with each of these statements (on a scale of 1 - strongly disagree to 4 - strongly agree) for respondents aged 10 to 15 is presented in the table below. We can see that swimming is correlated with agreement with the statements indicating confidence and disagreement with the statements indicating lack of confidence - a rather consistent pattern of a positive relationship between swimming and individual development (self-confidence / self-efficacy).

Table 5. Association between swimming and self-confidence in USoc Youth

Outcome variable	USoc Youth - Age 10 to 15	TP
	OLS	FE
I feel I have a number of good qualities	0.057***	0.034*
I don't have much to be proud of	-0.056***	-0.015
I certainly feel useless at times	-0.030*	-0.016
I am as able as most people	0.034***	0.029
I am a likeable person	0.037***	0.022
I can usually solve my own problems	0.028**	0.040*
I am inclined to feel that I am a failure	-0.043***	0.039
At times I feel I am no good at all	-0.032**	0.012

*Notes: Model, sample and dataset specified in column header. Coefficients of 'swimming and diving in the last 12 months' variable shown. All models include control variables. List of control variables provided in Methodology Section. Stars indicate statistical significance levels: *** < 1%; ** < 5%; * < 10% significance. Heteroscedasticity-robust standard errors used.*

Indoor and outdoor swimming

In the table below we can see the differences in the association between swimming and the outcomes in the TP data if we also make the distinction between indoor and outdoor swimming. Note that the different types of swimming are included in the regression model simultaneously, which means that the coefficients of indoor and outdoor swimming represent any incremental association over and above that of swimming in general.

What we can see from this analysis is that for a majority of the wellbeing outcomes it does not matter if one does indoor or outdoor swimming - the positive relationship with the outcome will be roughly the same. This is indicated by the fact that the incremental association with these outcomes of indoor and outdoor swimming is either statistically insignificant or very close to each other. Such is the case for life satisfaction, sense of a worthwhile life, trust, and meeting friends. However, outdoor swimming seems to be associated with significantly higher increases in happiness and general health, and also somewhat higher volunteering. Indoor swimming is associated with a stronger sense of belonging to the local area.

Table 6A. OLS results - Association between general, indoor and outdoor swimming and wellbeing/social outcomes: TP

Outcome variable	TP		
	OLS		
	Any swimming	Indoor swimming	Outdoor swimming
Life satisfaction (1 to 7 or 0 to 10)	0.094**	-0.021	0.024
Happiness (0 to 10)	0.019	0.043*	0.089***
Anxiety (0 to 10)	-0.050	0.080	0.032
Worthwhile life (0 to 10)	0.110**	0.002	-0.038
Trust people in general (1 to 3)	0.036	-0.006	-0.016
Trust people in neighbourhood (1 to 5)	0.057**	0.011	0.027
Frequency of meeting friends (1 to 5)	0.108***	0.025	-0.008
Feel belonging to local area (1 to 4)	-0.036	0.073***	0.010
People from different backgrounds get on well in local area (1 to 4)	-0.021	0.029	0.016
Spend time with family (0/1)	0.038***	0.009**	0.028***
Volunteered in the last 12 months (0/1)	0.006	0.048***	0.031***
Volunteered in sport in the last 12 months (0/1)	-0.008*	0.025***	0.020***
General health (1 to 5)	0.108***	-0.002	0.043***

*Notes: Model, dataset and main explanatory variable specified in column header. 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 Methodology Section. Stars indicate statistical significance levels: *** < 1%; ** < 5%; * < 10% significance. Heteroscedasticity-robust standard errors used.*

Swimming frequency

We then proceed to investigate the differences between more and less frequent swimming in terms of wellbeing uplifts. The TP and AL datasets provide information on how many sessions of swimming the respondents performed in the past 28 days. This allows us to construct an approximate indicator of swimming frequency.

The findings present a rather clear upward pattern showing that more frequent swimming is associated with a higher wellbeing uplift when compared to less frequent swimming. The positive wellbeing association of swimming at least once a week is clearly higher that than of swimming one to three times a month.

Table 6B. OLS results - Association between swimming frequency and life satisfaction: TP and AL

Swimming frequency in the past 28 days	TP	AL
No swimming (reference group)	0.000	0.000
Once a month	0.053	0.142***
Two-three times a month	0.029	0.183***
Once to twice a week	0.105***	0.245***
Twice a week or more	0.185***	0.281***

Disaggregated analysis

The 2018 HM Treasury Green Book identifies the importance of the 'distributional impact' of a policy, or how it affects different groups in society. To this end, this subsection investigates whether the association between swimming and the social outcomes in question is stronger for particular groups of the population. This is achieved through disaggregated regression analysis where the swimming variable is interacted with various demographic factors of interest. This gives evidence on how the positive association between swimming and wellbeing varies across different demographic groups.

Understanding Society (USoc)

Key findings.

- The positive association between health and swimming increases considerably as people get older. Older respondents also show a stronger association between swimming and community development.
- Younger respondents, in turn, have a stronger correlation of swimming with life satisfaction.
- Women have a stronger positive association with swimming across all outcomes except general health.
- Higher SEG have slightly higher positive associations of swimming with almost all relevant outcomes. However, the fact that the coefficients are positive and significant for all socioeconomic groups suggests that swimming does not have a significant distributional impact (i.e. it is not more beneficial for particular SEG).
- Scotland, Wales and Northern Ireland seem to have lower mental wellbeing associations with swimming than England. However, the associations with other outcomes are in some cases greater.
- Urban respondents have stronger correlations of swimming with health, wellbeing and relying on friends.

Table 7A. Heterogeneous effects analysis in the USoc data, waves 2 and 5

Sporting Futures Outcome Area	Physical wellbeing	Mental wellbeing	Individual development	Social and community development	
	General health (1 to 5)	Life satisfaction (1 to 7)	Can achieve goals (1 to 4)	Can rely on friends (1 to 4)	Volunteering (0/1)
SEG (based on occupation)[15]					
Higher	0.153***	0.087***	0.038***	0.078***	0.044***
Lower	0.068***	0.049***	-0.013	0.059***	0.031***
Other[16]	0.244***	0.106***	0.037***	0.065***	0.066***
SEG (income proxy)					
Income at or above median	0.168***	0.086***	0.021***	0.079***	0.039***
Income below median	0.133***	0.072***	0.018*	0.053***	0.057***
Age category					
16-34	0.071***	0.107***	0.035***	0.063***	0.039***
35-64	0.145***	0.064***	0.009	0.062***	0.039***
65+	0.418***	0.082***	0.032*	0.108***	0.116***
Gender					
Female	0.137***	0.083***	0.029***	0.081***	0.049***
Male	0.175***	0.077***	0.008	0.049***	0.044***
Region (grouped)					
North of England	0.180***	0.099***	0.023	0.066***	0.036***
Midlands	0.139***	0.101***	0.027*	0.056***	0.056***
South and East England	0.147***	0.075***	0.011	0.075***	0.045***
London	0.159***	0.082**	0.018	0.074***	0.060***

15 We suggest focus on the occupational lower SEG measure. The income-based measure is not available in AL.
16 Including students, retired people, and those not in the labour force for other reasons.

Wales	0.207***	0.065*	0.047*	0.052**	0.037***
Scotland	0.140***	0.052	-0.016	0.084***	0.040***
Northern Ireland	0.068**	0.041	0.074***	0.044*	0.072***
Urbanisation					
Urban	0.160***	0.090***	0.019**	0.076***	0.048***
Rural	0.130***	0.051**	0.023*	0.043***	0.045***

*Notes: OLS regressions. Coefficients of swimming 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 described in the Methodology Section. Stars indicate statistical significance levels: *** < 1%; ** < 5%; * < 10% significance. Heteroskedasticity-robust standard errors used.*

Understanding Society Youth (USoc Youth) (aged 10-15)

Key findings.

- Girls have considerably higher increases in wellbeing, health and self-confidence associated with swimming, compared to boys.
- Teenagers have a stronger correlation with life satisfaction, whereas younger children have a stronger correlation with solving own problems and the number of close friends.
- Respondents from lower SEG families have stronger associations of swimming with individual and community development outcomes (but weaker with life satisfaction and health).
- Children and teenagers from London do not show a statistically significant relationship between swimming and life satisfaction or health, but the relationship with self-efficacy is the strongest among all regions.

Table 7B. Heterogeneous effects analysis in the USoc Youth data, waves 2, 4, 6

Sporting Futures Outcome Area	Physical wellbeing	Mental Wellbeing	Individual development	Social and community development	
	General health (1 to 5)	Life satisfaction (1 to 7)	Can solve own problems (1 to 4)	Number of close friends	Volunteering (0/1)
SEG (income proxy)					
Income at or above median	0.147***	0.080***	0.023	0.126**	0.063***
Income below median	0.111***	0.040	0.032*	0.177***	0.073***
Age category					
10-12	0.138***	0.039	0.034**	0.163***	0.064***
13-15	0.120***	0.087***	0.020	0.135**	0.073***
Gender					
Male	0.088***	0.013	0.019	0.143**	0.089***
Female	0.172***	0.109***	0.036**	0.157***	0.047***
Region (grouped)					
North of England	0.147***	0.038	0.016	0.108	0.038*
Midlands	0.137***	0.018	0.055*	0.109	0.102***
South and East England	0.097***	0.097**	0.023	0.062	0.058***
London	0.047	0.022	0.071**	0.243**	0.084***
Wales	0.292***	0.132*	0.029	0.310**	0.036
Scotland	0.166***	0.067	-0.050	0.216*	0.094***
Northern Ireland	0.145***	0.082	0.033	0.282**	0.081**
Urbanisation					
Urban	0.127***	0.049**	0.031**	0.168***	0.065***
Rural	0.139***	0.097**	0.016	0.093	0.078***

*Notes: OLS regressions. Coefficients of swimming 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 described in the Methodology Section. Stars indicate statistical significance levels: *** < 1%; ** < 5%; * < 10% significance. Heteroskedasticity-robust standard errors used.*

Active Lives (AL)

Key findings.

- Higher SEG experience stronger associations with life satisfaction and health (reduction in the BMI), whereas lower SEG have a stronger correlation with volunteering.
- Younger swimmers experience the highest life satisfaction increase, whereas it is the middle-aged respondents this time that experience the strongest BMI reduction and trust increase.
- Women have a slightly stronger wellbeing association, but a lower BMI association.
- London has by far the strongest association with BMI reduction, and also with self-efficacy and trust (but the weakest with volunteering).

Table 7C. Heterogeneous effects analysis in the AL data, wave 2

Sporting Futures Outcome Area	Physical wellbeing	Mental wellbeing	Individual development	Social and community development	
	Body mass index	Life satisfaction (0 to 10)	Can achieve goals (1 to 5)	Trust people in local area (1 to 5)	Volunteering (0/1)
SEG (based on occupation)					
Higher	-0.408***	0.185***	0.069***	0.085***	0.061***
Lower	-0.289***	0.134**	0.069***	0.058**	0.105***
Other[17]	-0.045	0.299***	0.064**	-0.007	0.161***
Age category					
16-34	-0.118*	0.245***	0.085***	0.053***	0.072***
35-64	-0.498***	0.169***	0.068***	0.100***	0.078***
65+	-0.251***	0.173***	0.047***	0.016	0.064***
Gender					
Male	-0.425***	0.152***	0.072***	0.073***	0.086***
Female	-0.316***	0.217***	0.065***	0.076***	0.065***
Region (grouped)					
North of England	-0.330***	0.153***	0.063***	0.069***	0.074***
Midlands	-0.335***	0.251***	0.059***	0.043**	0.077***
South and East England	-0.325***	0.185***	0.070***	0.086***	0.079***
London	-0.655***	0.169***	0.089***	0.097***	0.049***
Urbanisation					
Urban	-0.353***	0.189***	0.070***	0.075***	0.073***
Rural	-0.406***	0.184***	0.063***	0.074***	0.078***

*Notes: OLS regressions. Coefficients of swimming 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 described in the Methodology Section. Stars indicate statistical significance levels: *** < 1%; ** < 5%; * < 10% significance. Heteroskedasticity-robust standard errors used.*

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

Active Lives Children and Young People (ALY) (aged 5-16)

Key findings.

- Confirming the USoc Youth findings, girls have considerably higher increases in wellbeing and self-efficacy and trust associated with swimming, compared to boys.
- The positive association with all outcomes goes up with age.
- Respondents from lower SEG families have stronger associations of swimming with all outcomes (but negligible difference for trust).
- Children and teenagers from London show a weaker relationship between swimming and happiness and do not have a statistically significant association with trust (the latter is also the case for the North of England).

Table 7D. Heterogeneous effects analysis in the ALY data, wave 1

Sporting Futures Outcome Area	Mental Wellbeing		Individual development	Social and community development	
	Happiness (0 to 10)	Life satisfaction (0 to 10)[18]	Can achieve goals (1 to 5)	Trust people in local area (1 to 5)	Volunteering (0/1)
SEG					
Lower	0.451***	0.375***	0.082***	0.053**	0.087***
Medium	0.285***	0.267***	0.073***	0.034***	0.069***
Higher	0.216***	0.324***	0.069***	0.052***	0.049***
Age category					
5-7	0.256***		0.049*	-0.022	0.330
8-10	0.258***		0.047***	0.039***	0.042***
11-13	0.333***	0.283***	0.094***	0.050***	0.050***
14-16	0.351***	0.326***	0.132***	0.066**	0.129***
Gender					
Male	0.148***	0.175**	0.053***	0.032***	0.061***
Female	0.436***	0.430***	0.092***	0.051***	0.069***
Region (grouped)					
North of England	0.357***	0.370***	0.094***	0.024	0.070***
Midlands	0.326***	0.282***	0.056***	0.039**	0.088***
South and East England	0.265***	0.261***	0.060***	0.060***	0.055***
London	0.221***	0.316***	0.093***	0.030	0.059***
Urbanisation					
Urban	0.309***	0.303***	0.073***	0.039***	0.065***
Rural	0.211***	0.275***	0.079***	0.058***	0.069***

*Notes: OLS regressions. Coefficients of swimming 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 described in the Methodology Section. Stars indicate statistical significance levels: *** < 1%; ** < 5%; * < 10% significance. Heteroskedasticity-robust standard errors used.*

18 Only available for respondents in school year 7 or higher.

Taking Part (TP)

Key findings.

- Only life satisfaction is consistently more strongly associated with swimming for higher economic groups. The remaining numbers are ambiguous and the situation changes when we switch from job-based to income-based socio-economic definitions.
- The upward trend of health increases as the respondents get older is confirmed. This time we also see a similar trend for life satisfaction.
- There is a higher association of swimming with wellbeing for women.
- There is also a geographic pattern in this dataset: wellbeing and health correlations increase as we move north.

Table 7E. Heterogeneous effects analysis in the TP data, waves 1-3 + 6-12

Sporting Futures Outcome Area	Physical wellbeing	Mental wellbeing	Social and community development	
	General health (1 to 5)	Life satisfaction (0 to 10)	Trust neighbours (1 to 4)	Volunteering (0/1)
SEG (based on occupation)				
Higher	0.117***	0.088***	0.079***	0.050***
Lower	0.145***	0.046	0.062***	0.060***
Other[19]	0.068***	0.015	0.132***	0.104***
SEG (income proxy)				
Income at or above median	0.128***	0.076***	0.068***	0.044***
Income below median	0.114***	0.055	0.092***	0.084***
Age category				
16-34	0.042***	-0.016	0.062***	0.048***
35-64	0.109***	0.079***	0.088***	0.055***
65+	0.348***	0.152***	0.059**	0.087***
Gender				
Female	0.119***	0.089***	0.080***	0.059***
Male	0.130***	0.046*	0.072***	0.056***
Region (grouped)				
North of England	0.148***	0.129***	0.091***	0.048***
Midlands	0.129***	0.094**	0.089***	0.051***
South and East England	0.110***	0.019	0.052***	0.060***
London	0.092***	0.028	0.091***	0.083***
Urbanisation				
Urban	0.119***	0.077***	0.073***	0.055***
Rural	0.146***	0.045	0.050**	0.070***

*Notes: OLS regressions. Coefficients of swimming 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 described in the Methodology Section. Stars indicate statistical significance levels: *** < 1%; ** < 5%; * < 10% significance. Heteroskedasticity-robust standard errors used.*

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

Detailed age and gender breakdown

We continue with a more detailed breakdown of the wellbeing association by age and gender. This analysis is performed for life satisfaction (Table 7F) and health (Table 7G).

Table 7F. Heterogeneous effects of swimming on life satisfaction (0 to 10): detailed age and gender breakdown

Dataset	USoc (1 to 7)	AL	TP
Swimming female			
16-24	0.110***	0.148**	-0.015
25-34	0.107***	0.223***	-0.006
35-44	0.095***	0.220***	0.106**
45-54	0.028	0.221***	0.118**
55-64	0.072**	0.261***	0.076
65-74	0.100**	0.130**	0.205***
75+	0.045	0.544***	0.146
Swimming male			
16-24	0.144***	0.600***	0.118
25-34	0.086***	0.198***	-0.109*
35-44	0.080***	0.013	0.043
45-54	-0.011	0.024	0.032
55-64	0.120***	0.207***	0.081
65-74	0.102**	0.187***	0.127**
75+	-0.071	0.160	0.054

Table 7G. Heterogeneous effects of swimming on general health (1 to 5): detailed age and gender breakdown

Dataset	USoc	TP
Swimming female		
16-24	-0.079***	-0.034**
25-34	0.064***	0.030**
35-44	0.116***	0.070***
45-54	0.133***	0.118***
55-64	0.241***	0.195***
65-74	0.460***	0.326***
75+	0.644***	0.517***
Swimming male		
16-24	0.162***	0.153***
25-34	0.140***	0.061***
35-44	0.107***	0.053***
45-54	0.136***	0.089***
55-64	0.186***	0.136***
65-74	0.328***	0.288***
75+	0.419***	0.407***

*Notes: OLS regressions. Coefficients of swimming 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 described in the Methodology Section. Stars indicate statistical significance levels: *** < 1%; ** < 5%; * < 10% significance. Heteroskedasticity-robust standard errors used.*

Analysis for life satisfaction split by age and gender reveals that the wellbeing uplift is the most consistent for both genders between the ages of 65-74. Females aged 35-44 (the phase where they are most likely to be mothers with young children) also experience a strong positive association.

While swimming benefits the wellbeing and general health of people of all ages, the positive health correlation with swimming increases as people get older.

Results summary

Consistently across different datasets we see a positive association between swimming and the following outcomes:

- social connections - having friends, number of close friends, relying on friends, spending time with friends and family
- trust - trusting people in general, trusting neighbours
- community cohesion - belonging to local area
- volunteering - in general and in sport (all datasets) also with frequency of volunteering
- perceived ability to achieve goals (USoc and AL)
- life satisfaction (all datasets) and happiness (wherever measured), worthwhile life (wherever measured) and anxiety (AL but not TP)
- health (all datasets), mental health (USoc) and BMI (AL)
- for young people we can see positive co-movement with wellbeing, health, self-efficacy, trust, volunteering and social connections.

Although the data structure does not have a good enough panel component to allow for a robust FE analysis, it is encouraging to see that some relationships still maintain their statistical significance in a FE model (in particular for those outcomes in the TP data that are collected over a longer time series).

Overall, the analysis paints a picture of the average swimmer as happier, healthier, more socially engaged and involved in the life of their community.

In particular, outdoor swimming has a particularly strong correlation with happiness and health, whereas indoor swimming has a stronger relationship with community cohesion and volunteering. Furthermore, swimming more often is associated with higher wellbeing improvements than swimming less often, with particularly consistent increases observed for swimming once a week or more.

Disaggregated analysis shows a few trends that emerge consistently across datasets. First, older respondents are the ones where swimmers have the highest health differential compared to non-swimmers. Second, women experience higher increases in life satisfaction and self-confidence associated with swimming - this is especially true for younger girls. In terms of SEC, higher SEG generally benefit more from swimming, although there are exceptions to the rule.

Annex 1: Detailed methodology - technical description

Datasets used

We use five nationally representative datasets to investigate the relationship between wellbeing / social capital and swimming.

- The **Understanding Society (USoc)** dataset is the successor of the discontinued British Household Panel Survey (BHPS) based at the Institute for Social and Economic Research at the University of Essex. 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. (UK-wide data).
- The **Understanding Society Youth (USoc Youth)** dataset is collected alongside the main USoc dataset using a separate questionnaire designed for young people aged 10-15. It collects a more limited, but still relevant set of demographics and social outcomes, and participation in sport activities such as swimming. (UK-wide data).
- The **Taking Part (TP)** dataset is a nationally representative database commissioned annually by the Department for Digital, 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, as well as the usual socio-demographic information on respondents. (England data).
- The **Active Lives (AL)** dataset is a Sport England-led survey about people in England (aged 14+) and their participation in leisure and recreational activities, including numerous kinds of sport, physical activity and culture. Wellbeing information is also collected. The survey has an impressive sample, large enough to be representative for the adult population in each local authority area in England. (England data).
- The **Active Lives Children and Young People (ALY)** survey was carried out by Sport England during the 2017-18 school year. The target population are schoolchildren aged 5-16 (corresponding to years 1-11 of schooling[20]). The collected data broadly follows the AL format, but with some modifications - the main focus is on the respondents' involvement in sport activity, where a lot of detail is collected. The sample size is likewise impressive. (England data).

Sample sizes

The key variables of interest for the study - swimming 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:

Table B1. Datasets used for this study

Dataset	Years (waves) Used	Total sample[21]	Of which are swimmers
USoc	2010/11, 2013/14 (waves 2, 5)	90194	29287
USoc Youth	2010/11, 2012/13, 2014/15 (waves 2, 4, 6)	12320	4676
TP	2005/06 - 2007/08, 2010/11 - 2016/17 (waves 1-3, 6-12)	150834	51944
AL	2016/17 (wave 2)	107469	34380
ALY	2017/18 (wave 1)	108325	24150

20 But only respondents from year 3 and above were asked about relevant outcomes, whereas life satisfaction is only available from school year 7.
21 Number of respondents for whom swimming information is available.

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.

- **Swimming**

Swimming or diving in the last 12 months (with slight variations across datasets) is the main 'treatment' variable of this study, that is, the main variable whose benefits we are trying to identify. This swimming frequency - at least once in the last 12 months - was chosen because it is the most readily available frequency in national datasets.

Table C1. Swimming variables by dataset

Dataset	Years used	Swimming variables	Response scale
USoc	Wave 2 (2010/11) and 5 (2013/14)	Swimming or diving in the last 12 months	Binary (yes/no)
USoc Youth	Waves 2, 4, 6 (2010/11, 2012/13, 2014/15)	Does swimming or diving (timeframe not specified)	Binary (yes/no)
TaP	Waves 1-12 (2005/06 - 2016/17)	Swimming or diving in the last 12 months [indoors] Swimming or diving in the last 12 months [outdoors] Swimming or diving in the last 12 months [any] - true if any of the above are true	All binary (yes/no)
AL	Wave 2 (2016/17)	Swimming in the last 12 months (incl. diving, water polo and water-based class)	Binary (yes/no)
ALY	Wave 1 (2017/18)	Minutes spent swimming and diving in the last 7 days (recoded into binary showing whether > 0) Ability to swim 25m unaided (supplementary analysis)	Binary (yes/no)

- **Physical wellbeing and Mental wellbeing outcomes**

There are a range of variables which measure the respondent's subjective wellbeing and general health. Among these are the four 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 C2. Physical development and Mental wellbeing 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 (only from wave 9 in TP)
Happiness	How happy did you feel yesterday?	0 to 10	TP, AL, ALY
Anxiety	How anxious did you feel yesterday?	0 to 10	TP (wave 9 onwards), AL
Worthwhile life	To what extent do you feel the things you do in your life are worthwhile?	0 to 10	TP (wave 9 onwards), AL, ALY
General health	In general, would you say your health is... (multiple choice)	1 (poor) to 5 (excellent)	All except AL and ALY
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	USoc
BMI	Derived from the respondent's self-reported weight and height: weight divided by height squared.	Continuous (in kg/m2), but categorical version also available, ranging from 1 (underweight) to 5 (morbidly obese)	AL

- **Individual development and Social and community development outcomes**

The aim of this study is to gather evidence in support of a benefit of swimming on individual development/ self-efficacy, as well as social and community development outcomes. The datasets contain a range of variables that fit within these outcome areas. These variables cover topics such as being able to achieve goals, trust, volunteering, number and quality of friendships, affinity to one's community, and closeness with one's family. A list of the key outcomes present in different datasets and analysed in this study can be seen below:

Table C3. Individual development and Social and community development outcomes by dataset

Wellbeing variable	Question form	Response scale (after recoding)*	Datasets used in
Self-efficacy	To what extent do the following set of statements describe you? - It is easy for me to stick to my aims and accomplish my goals.	1 to 4 (USoc and AL Youth) , 1 to 5 (AL)	USoc, AL, ALY
Self-confidence / self-efficacy	8 statements with which the respondent can agree or disagree (e.g. "I have a number of good qualities").	1 to 4	USoc Youth
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 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
Satisfaction with friends	How do you feel about your friends? (Youth respondent is presented with 7 faces ranging from sad to happy).	1 to 7	USoc Youth
Spend time with family	Do you spend time with family/friends? (TP) How many times in the last 7 days have you eaten evening meals with your family? (USoc Youth).	Binary (TP), 1 to 4 (USoc Youth)	TP, Usoc Youth
Frequency of meeting friends	How often do you meet up with friends?	1 (Never) to 5 (Most days)	TP
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	TP (until wave 7 only)
Trust in people in neighbourhood	Do you believe most people in your neighbourhood / local area can be trusted? (Varies slightly by dataset).	1 to 4 (TP, AL Youth) 1 to 5 (Active Lives)	TP (until wave 7 only), AL, ALY
Belonging to neighbourhood	How strongly do you feel you belong to your local area?	1 to 4	TP
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 to 4	TP
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)	USoc, USoc Youth, TP
Frequency of volunteering	Approximately how many hours did you spend volunteering in the last 4 weeks? (USoc). How often do you do voluntary or community work? (Usoc Youth).	Continuous (USoc), 1 to 6 (USoc Youth)	USoc, USoc Youth
Volunteering in sport	Was this [volunteering] connected to any of the following areas? - Sport (TP). Volunteered in the last 12 months to support sport or physical activity (AL, ALY). Volunteered in sport more than once in the last 12 months (AL, ALY). Volunteered in sport in the last 4 weeks (AL).	All binary (yes/no)	TP, AL, ALY

**Wherever appropriate, the scales have been inverted so that a higher value indicates a more desirable/positive outcome.*

- **Demographic control variables**

There are a range of factors that are known to have a great influence on our subjective wellbeing. By including these variables in the analysis we can control for these determinants of our wellbeing, so that we can see just the change in wellbeing attributed to swimming.

It is a consistent finding across datasets that sports participation and sport group membership in general, and therefore implicitly also swimming in particular, is more common amongst higher SEG in society. At the same time, the more money you have, the more likely you are (on average) to have better health, more happiness and overall life satisfaction. Affluence and earnings are positively associated with the five outcomes from the Government's ***Sporting Future***[22] strategy. It is therefore very important to control for (effectively cancel out the effect of) demographic characteristics in order to avoid biased estimates of the impact of sport (and swimming in particular) on wellbeing.

Fujiwara and Campbell (2011)[23] draw up a list of 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 swimmers and non-swimmers comparatively, and to paint a demographic profile of the average swimmer. A list of demographic variables used in this study and their availability by dataset can be found below:

Table C4. Demographic control variables by dataset

Wellbeing variable	Response scale / categories	Notes / Comments	Availability by dataset				
			USoc	USoc Youth	TP	AL	ALY
Age	Whole numbers indicating age in years; 5-year bands	TP stopped collecting exact age in wave 12 and only collects age bands.	V	V	V	V	V
Gender	Male, Female		V	V	V	V	V
Income	Continuous (natural logarithm)	USoc has exact income. TP has income bins, where we impute income as the midpoint. We then take the natural logarithm in all cases. TP has personal income; USoc has household income. AL has no income data.	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	
Number of children in household	0, 1, 2, 3, 4+	Four or more children grouped together to avoid small bin sizes. Usually defined as children aged 0-15 in the household, but USoc Youth uses 0-13 and some datasets don't specify the age.	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. AL bundles all level 4 education (Certificate of HE and above) together.	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	

22 https://www.gov.uk/government/publications/sporting-future-a-new-strategy-for-an-active-nation

23 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.

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 USoc Youth because it is not collected in the same wave or questionnaire group as the outcomes.	V		V		
Ethnicity (broad categories)	White, Mixed, Asian, Black, other	In USoc Youth not collected in the same wave as swimming.	V		V	V	V
Disability	No disability, non-limiting disability, limiting disability.	AL only.				V	V
Socio-economic class (job-based)	Higher SEC, lower SEC, SEC residuals	In most datasets, this is based on the 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).			V	V	
Socio-economic 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		
Socio-economic class (possession-based)	Low, medium, high	The ALY version of the SEC classification is also known as the FAS and is based on the answers to a series of questions about home possessions and circumstances.					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				
Wave of survey	Dataset-dependent	Indicator for each wave in the dataset. Included to account for time trends.	V	V	V		
Interview month	January to December	Included to account for seasonality.	V	V	V	V	
School term	Spring, summer, autumn	Included to account for seasonality in ALY.					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 (SEC)

SEC is an important driver of wellbeing which we control for in our analysis[24]. Furthermore, in the disaggregated analysis, where we look at how results vary for different groups in society, we look at each outcome and swimming participation separately for higher and lower SEG. For a better understanding of these results, we explain the criteria used to differentiate these SEG.

24 We also control for wider 'household income' in the data where possible, so we can be confident we are capturing socio-economic characteristics in the analysis.

There are three measures of SEC 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
- the third is used in the ALY data and is based on information about household possessions and spending.

The first of the two measures is divided into THREE SEG 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 which includes students, retired, people with disabilities, stay-at-home people and others.

The second measure is divided into two categories:

- higher SEG are those respondents whose household income is at or above the sample median
- lower SEG are those respondents whose household income is below the median.

The third measure is divided into three ordered categories - low, medium and high, based on a recoding of the raw score of 0-20 that results from the answers to the survey questions about household possessions and spending.

The model

The analysis starts off by tabulating wellbeing and social outcomes as well as demographics for swimmers and non-swimmers. While these tables may suggest the outcomes and swimming have a positive association, we do not know for sure whether this indicates that swimming leads to a wellbeing or social capital benefit, or if something else is responsible.

This leads us to move on to regression analysis, where we are able to account for the many determinants of wellbeing/social outcomes that we listed in the "Demographic control variables" section and enable us to isolate the effect of swimming on these outcomes. This is achieved by including them in an Ordinary Least Squares (OLS) regression equation such as:

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

Here, O_i is an outcome variable selected from the lists mentioned in the subsections of "Physical wellbeing and Mental wellbeing outcomes" and "Individual development and Social and community development outcome"; a separate regression is run for each outcome variable in each dataset. Sw_i is a dummy variable indicating whether the respondent has done any swimming or diving in the last 12 months. X_i is a vector of all the control variables available in the dataset, as listed in the "Demographic control variables" subsection; and ε is the error term containing unobserved factors that determine the relevant outcomes[25].

To further insure against the influence of unobserved factors (a.k.a. omitted variable bias), we supplement OLS analysis by FE regressions (in all datasets except for AL, where a panel component is not available). 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. FE generally confirm OLS results but mostly lack statistical significance. This is due to the poor panel structure of the data - either because our key variables are not collected in every wave (USoc, TP) or because most individuals are not tracked over time and new respondents are mostly recruited each wave (TP).

As the panel data structure is not good enough to conduct robust panel data analysis, we fall back to OLS results as our main findings and supplement these with the caveat that the regression coefficient only represents an association between swimming and the wellbeing/social outcome, holding a range

25 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 swimming 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 swimming at the same time.

of demographic factors constant, but does not necessarily indicate a causal effect or a direct benefit of swimming on wellbeing / social capital.

Furthermore, we can investigate how the association between outcomes and swimming varies by age category, gender, SEC and other demographic variables of interest. This is done with the help of regression models with interaction terms, such as Equation 2 below (where SEG_i is a SEG dummy). This would feed into the objective of identifying whether swimming is associated with higher improvements in social outcomes or wellbeing for more socially vulnerable groups.

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

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 contexts[26], and at the same time it facilitates interpretation and even subsequent monetary valuation of the results.

26 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

01996SE-2 10/19